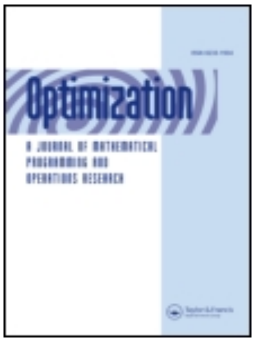

Taylor & Francis
Taylor & Francis Group

Optimization

A Journal of Mathematical Programming and Operations Research

ISSN: 0233-1934 (Print) 1029-4945 (Online) Journal homepage: http://www.tandfonline.com/loi/gopt20

Interior proximal bundle algorithm with variable metric for nonsmooth convex symmetric cone programming

Felipe Alvarez & Julio López

To cite this article: Felipe Alvarez & Julio López (2016) Interior proximal bundle algorithm with variable metric for nonsmooth convex symmetric cone programming, Optimization, 65:9, 1757-1779, DOI: 10.1080/02331934.2016.1193863

To link to this article: http://dx.doi.org/10.1080/02331934.2016.1193863

Published online: 07 Jun 2016.

OPTIMIZATION, 2016
VOL. 65, NO. 9, 1757–1779
http://dx.doi.org/10.1080/02331934.2016.1193863

Taylor & Francis
Taylor & Francis Group

Interior proximal bundle algorithm with variable metric for nonsmooth convex symmetric cone programming

Felipe Alvarez[a] and Julio López[b]

[a]Departamento de Ingeniería Matemática, Centro de Modelamiento Matemático (CNRS UMI 2807), FCFM, Universidad de Chile, Santiago, Chile; [b]Facultad de Ingeniería, Universidad Diego Portales, Santiago, Chile

ABSTRACT
This paper is devoted to the study of a bundle proximal-type algorithm for solving the problem of minimizing a nonsmooth closed proper convex function subject to symmetric cone constraints, which include the positive orthant in $\mathbb{R}^n$, the second-order cone, and the cone of positive semidefinite symmetric matrices. On the one hand, the algorithm extends the proximal algorithm with variable metric described by Alvarez et al. to our setting. We show that the sequence generated by the proposed algorithm belongs to the interior of the feasible set by an appropriate choice of a regularization parameter. Also, it is proven that each limit point of the sequence generated by the algorithm solves the problem. On the other hand, we provide a natural extension of bundle methods for nonsmooth symmetric cone programs. We implement and test numerically our bundle algorithm with some instances of Euclidean Jordan algebras.

ARTICLE HISTORY
Received 24 November 2015
Accepted 19 May 2016

KEYWORDS
Euclidean Jordan algebra; symmetric cone programming; variable metric; proximal method; bundle methods

1. Introduction

Let $\mathbb{V} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ be a Euclidean Jordan algebra (cf. Section 2) and consider the following symmetric cone programming (SCP) problem:

$$\text{(SCP)} \qquad f_* = \min f(x) \text{ s:t: } w(x) = \mathbb{A}x + b \in \mathcal{K},$$

where $f : \mathbb{R}^n \to (-\infty, +\infty]$ is closed proper convex, $\mathbb{A} : \mathbb{R}^n \to \mathbb{V}$ is a linear application defined by $\mathbb{A}x = \sum_{k=1}^{n} x_k a_k$, where $a_k \in \mathbb{V}$ for $k = 1, \ldots, n$, $b \in \mathbb{V}$, and

$$\mathcal{K} := \{x \circ x : x \in \mathbb{V}\}$$

denotes the *cone of square elements* in $\mathbb{V}$. Such a formalism covers several cones as the positive orthant in $\mathbb{R}^n$, the second-order cone, and the cone of positive semidefinite symmetric matrices. SCP thus provides a unified framework for several classes of optimization problems such as nonlinear programming (NLP), second-order cone programming (SOCP) and semidefinite programming (SDP), and hence has extensive applications in engineering, economics, game theory, management science, and other fields; see [2–6] and references therein.

In recent years, instances of SCP with linear and quadratic objective functions have attracted the attention of some researchers for the development of interior-point methods similar to those that exist for linear programming (see e.g. [7–12]) and quadratic programming (see e.g. [13–15]). For instance, Gu et al. [9] generalized Roos et al.'s algorithm for linear optimization (LO) in [16] to linear

CONTACT Julio López julio.lopez@udp.cl

SCP, Wang et al. [17], Wang and Bai [12] extended Darvay's algorithm for LO in [18] to linear SDP and SCP, respectively.

In this paper, we focus on the case where the objective function f in SCP is convex and nonsmooth. Our approach consists in extending to convex SCP problems the *interior proximal algorithm with variable metric* (IPAVM) that was proposed in [1] for convex SOCP. The standard proximal algorithm (PA) was first introduced by Martinet [19] based on previous work by Moreau [20] and it was then further developed and studied by Rockafellar [21]. Later, several authors [22–25] generalized the PA for convex programming with nonnegative constraints, replacing the quadratic proximal term by a generalized distance-like function, which also plays the role of a sort of barrier that force the iterates to stay in the interior of the feasible set (strictly feasible solutions). Following a different approach, the IPAVM for SOCP developed in [1] uses a quadratic variable metric induced by a class of positive definite matrices together with a regularization parameter appropriately chosen so that the iterates be interior points. For the definition of the variable metric in the general case of SCP, the idea developed in this paper is to replace the positive definite matrix with a positive definite operator defined on $\mathbb{V}$, by specifically using the inverse of its quadratic representation as defined in Section 2. This allows us to propose a generic IPAVM for SCP and extend to it the theoretical analysis of [1].

We say that IPAVM is a generic algorithm in the sense that, at each iteration, it requires to solve an auxiliary regularized strongly convex minimization problem, whose numerical resolution would depend on the specific instance one is dealing with. As an illustration, the computational implementations and numerical experiences in [1] were developed only for smooth objective functions and performed for special instances of Linear SOCP, that is, when the objective function f is linear. Thus, although most theoretical results in [1] do not require differentiability on the objective function, the actual IPAVM that was implemented there was not adapted to solve a nonsmooth convex SOCP.

The auxiliary inner problem in IPAVM may be very hard to solve when the objective function is nonsmooth. Nonsmooth optimization problems are in general difficult to solve, even when they are unconstrained. They arise in many fields of applications, for example, in economics, engineering and optimal control. Among algorithms for nonsmooth optimization, we mention the subgradient,[26] cutting planes,[27] analytic center cutting-planes [28] and *proximal bundle methods*.[29–31] The latter class, which replaces the objective function with a polyhedral model involving accumulated subgradients, is the most robust and reliable algorithm because they not only stabilize the optimization procedure but make the subproblem a well-posed one, that is, with unique solution.

For unconstrained nonsmooth problems, iterates of a proximal bundle algorithm are generated by solving an auxiliary quadratic programming (QP) problem at each iteration. Each QP problem is defined by means of a cutting-planes model of the objective function, stabilized by a quadratic term centred at the best point obtained so far (which is referred to as the serious step). An important feature of bundle methods is that the size of each QP problem can be controlled via the so-called aggregation techniques.[29, Ch. 9] On the other hand, constrained nonsmooth problems are more complex, and only a few practical methods can be found in the literature. Convex problems with easy constraints (such as bound or linear constraints) can be solved either by inserting the constraints directly into each QP problem or by projecting iterates onto the feasible set.[32,33] For convex problems with general constraints of nonnegativity, one possibility is to solve an equivalent unconstrained problem with an exact penalty objective function.[34,35] Another possibility is to replace the quadratic proximal term with a Bregman-type distance, which produces interior schemes.[36,37] Some other bundle-type approaches for this last class of constraints has been proposed in [38]. Recently, in [39] the authors present an inexact spectral bundle method for solving a convex quadratic symmetric conic programming problem, where at each iteration an eigenvalue minimization problem is solved inexactly.

In this work, we use the technique of proximal bundle methods to develop an implementable version of IPAVM. In fact, we present and develop in detail a computational implementation of a method that replaces the original objective function with a polyhedral model involving accumulated subgradients of the objective function. A particular technical difficulty that we are able to overcome here is to ensure that the inexact bundle subiterations are interior points (strictly feasible solutions). The resulting *interior proximal bundle algorithm with variable metric* (IPBAVM) is suited to address nonsmooth convex SCP problems, including SOCP and SDP as special cases. To the best of our knowledge, the bundle method had not been studied yet in the context of a general convex SCP.

The paper is organized as follows. Section 2 reviews some basic results on Euclidean Jordan algebras. In Section 3, we describe our algorithm with variable metric induced by a positive definite operator and we prove the convergence of our algorithm. In Section 4, we develop a detailed implementation of the bundle algorithm that solves our problem. This implementation is applied to some test problems which are described in Section 5. Indeed, these test problems cannot be solved with the version implemented in [1] of IPAVM for convex SOCP. Finally, concluding remarks are given in Section 6.

Notation

The following notation is used throughout this paper. For a closed proper convex function f and, for some $\varepsilon \geq 0$, $\partial_\varepsilon f(x) = \{p \in \mathbb{R}^n : f(x) + \langle p, z - x\rangle - \varepsilon \leq f(z),\ \forall z \in \mathbb{R}^n\}$ denotes its ε-subdifferential at x, $\partial f = \partial_0 f$ its subdifferential.[40]

2. Preliminaries on Euclidean Jordan algebras

In this subsection, we briefly describe some concepts, properties and results from Euclidean Jordan algebras that are needed in this paper and they have become important in the study of conic optimization; see, e.g. Schmieta and Alizadeh [11], Faraut and Korányi [41].

A *Euclidean Jordan algebra* (EJA) is a triple $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$, where $(\mathbb{V}, \langle\cdot,\cdot\rangle)$ is a finite-dimensional space over the real field $\mathbb{R}$ equipped with an inner product $\langle\cdot,\cdot\rangle$, and the product $(x, y) \mapsto x \circ y : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ is a bilinear mapping satisfying the following three conditions:

(i) $x \circ y = y \circ x$ for all $x, y \in \mathbb{V}$,
(ii) $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in \mathbb{V}$ where $x^2 = x \circ x$, and
(ii) $\langle x \circ y, z\rangle = \langle y, x \circ z\rangle$ for all $x, y, z \in \mathbb{V}$,

and there exists a (unique) unitary element $e \in \mathbb{V}$ such that $x \circ e = x$ for all $x \in \mathbb{V}$. Hence orth, we simply say that $\mathbb{V}$ is a EJA and $x \circ y$ is called the *Jordan product* of x and y.

In a EJA, $\mathbb{V}$ is known that the set of squares $\mathcal{K} = \{x^2 : x \in \mathbb{V}\}$ is a *symmetric cone* (see [41, Theorem III.2.1]). This means that $\mathcal{K}$ is a self-dual closed and convex cone with interior int $(\mathcal{K}) \neq \emptyset$, and for any two elements $x, y \in$ int $(\mathcal{K}\propto$, there exists an invertible linear transformation $\mathcal{T} : \mathbb{V} \to \mathbb{V}$ such that $\mathcal{T}\swarrow\mathcal{K}\propto = \mathcal{K}$ and $\mathcal{T}(x) = y$.

The *rank* of $\mathbb{V}$ is defined as $r = \max\{\deg(x) : x \in \mathbb{V}\}$, where $\deg(x)$ is the degree of $x \in \mathbb{V}$ given by $\deg(x) = \min\{k > 0 : \{e, x, x^2, \dots, x^k\}$ is linearly dependent$\}$.

An element $c \in \mathbb{V}$ is an *idempotent* iff $c^2 = c$; it is a *primitive idempotent* iff it is nonzero and cannot be written as a sum of two nonzero idempotents. We say that a finite set $\{e_1, \dots, e_r\}$ of primitive idempotents in $\mathbb{V}$ is a *Jordan frame* iff $e_i \circ e_j = 0$ for all $i \neq j$, and $\sum_{i=1}^r e_i = e$. Note that $\langle e_i, e_j\rangle = \langle e_i \circ e_j, e\rangle = 0$ whenever $i \neq j$.

The following theorem gives us a spectral decomposition for the elements in a EJA (see Theorem III.1.2 of [41]).

Theorem 2.1: [Spectral decomposition theorem] *Suppose that $(\mathbb{V}, \circ, \langle\cdot,\cdot\rangle)$ is a EJA with rank r. Then, for every $x \in \mathbb{V}$, there exists a Jordan frame $\{e_1(x), \dots, e_r(x)\}$ and real numbers $\lambda_1(x), \dots, \lambda_r(x)$, arranged in the nonincreasing order, such that $x = \lambda_1(x)e_1(x) + \cdots + \lambda_r(x)e_r(x)$.*

The numbers $\lambda_i(x)$ (counting multiplicities), which are uniquely determined by x, are called the eigenvalues of x, and we write the maximum and the minimum eigenvalue of x as $\lambda_{\max}(x)$ and $\lambda_{\min}(x)$, respectively. The trace of x, denoted as $\mathrm{tr}(x)$, is defined by $\mathrm{tr}(x) := \sum_{j=1}^r \lambda_j(x)$.

It is easy to show that $x \in \mathcal{K}$ (resp. int $(\mathcal{K}\propto)$ iff every eigenvalue $\lambda_i(x)$ of x is nonnegative (resp. positive). Moreover, an element $x \in \mathbb{V}$ is *invertible*, if $\det(x) = \Pi_{j=1}^{r}\lambda_j(x) \neq 0$, that is, if every eigenvalue of x is nonzero.

Example 2.1: Typical examples of Euclidean Jordan algebras are the following:

(i) *Euclidean Jordan algebra of m-dimensional vectors:*

$$\mathbb{V} = \mathbb{R}^m,\ \mathcal{K} = \mathbb{R}^m_+,\ r = m,\ \langle x, y\rangle = \sum_{i=1}^{m} x_i y_i,\ x \circ y = (x_1y_1, \ldots, x_my_m).$$

Here, the unitary element is $e = (1, \ldots, 1) \in \mathbb{R}^m$, the spectral decomposition of any $x \in \mathbb{R}^m$ is $x = \sum_{i=1}^{m} x_i e_i$, where e_i denotes a vector with 1 in the i-th entry and 0's elsewhere, and $\lambda_{\min}(x) = \min(x)$, $\lambda_{\max}(x) = \max(x)$.

(ii) *Euclidean Jordan algebra of quadratic forms:* Let $\mathcal{L}^m_+ = \{x = (x_1, \bar{x}) \in \mathbb{R} \times \mathbb{R}^{m-1} : \|\bar{x}\| \leq x_1\}$ be the cone of squares, known as the second-order cone. Then,

$$\mathbb{V} = \mathbb{R}^m, \quad \mathcal{K} = \mathcal{L}^m_+,\ r = 2,\ \langle x, y\rangle = \sum_{i=1}^{m} x_i y_i, \quad x \circ y = (\langle x, y\rangle, x_1\bar{y} + y_1\bar{x}).$$

Here, the unitary element is $e = (1, 0, \ldots, 0) \in \mathbb{R}^m$, the spectral decomposition of any $x \in \mathbb{R}^m$ associated with $\mathcal{L}^m_+$ is given by $x = \lambda_1(x)u_1(x) + \lambda_2(x)u_2(x)$, where $\lambda_i(x) = x_1 + (-1)^i\|\bar{x}\|$ and $u_i(x) = \frac{1}{2}(1, (-1)^i\frac{\bar{x}}{\|\bar{x}\|})$, for $i = 1, 2$, which denote the eigenvalues and eigenvectors of x, respectively. Moreover, $\lambda_{\min}(x) = \lambda_1(x)$ and $\lambda_{\max}(x) = \lambda_2(x)$.

(iii) *Euclidean Jordan algebra of n-dimensional symmetric matrices:* Let $\mathcal{S}^n$ be the set of all $m \times m$ real symmetric matrices, and $\mathcal{S}^n_+$ be the cone of $m \times m$ symmetric positive semidefinite matrices. Then,

$$\mathbb{V} = \mathcal{S}^n, \quad \mathcal{K} = \mathcal{S}^n_+, \quad r = m, \quad \langle X, Y\rangle = \mathrm{tr}(XY), \quad X \circ Y = (XY + YX)/2.$$

Here tr denotes the trace of a matrix X. In this setting, the identity matrix $I \in \mathbb{R}^{m\times m}$ is the unit element, and the spectral decomposition of any $X \in \mathcal{S}^n$ is given by $X = \sum_{i=1}^{m} \lambda_i(X)q_i(X)q_i(X)^\top$, where $\lambda_i(X)$ and $q_i(X) \in \mathbb{R}^m$ denote the eigenvalue and eigenvector of X, respectively.

Other examples of EJA can be found in [41,42].

For any $a \in \mathbb{V}$, the *Lyapunov transformation* $L_a : \mathbb{V} \to \mathbb{V}$ and the *quadratic representation* $\mathcal{Q}_a : \mathbb{V} \to \mathbb{V}$ of a are defined as

$$L_a(x) := a \circ x \text{ and } \mathcal{Q}_a(x) := (2L_a^2 - L_{a^2})(x) = 2a \circ (a \circ x) - a^2 \circ x, \quad \forall x \in \mathbb{V}. \tag{1}$$

These transformations are linear and self-adjoint on $\mathbb{V}$ (see [41]). The quadratic representation is an essential concept in the theory of Jordan algebras and will play an important role in our subsequent development. In the following example, we describe these transformations in the EJAs defined in Example 2.1.

Example 2.2:

(i) For the EJA of m-dimensional vectors, the above transformations are: $L_a(x) = \mathrm{Diag}(a)x$ and $\mathcal{Q}_a(x) = \mathrm{Diag}(a^2)x$, where $\mathrm{Diag}(q)$ denotes a diagonal matrix of size m whose diagonal entries are the entries of q.

(ii) For the EJA of quadratic forms, one has that

$$L_a(x) = \begin{pmatrix} a_1 & \bar{a}^\top \\ \bar{a} & a_1 I \end{pmatrix} \begin{pmatrix} x_1 \\ \bar{x} \end{pmatrix}, \; \mathcal{Q}_a(x) = \begin{pmatrix} \|a\|^2 & 2a_1\bar{a}^\top \\ 2a_1\bar{a} & (a_1^2 - \|\bar{a}\|^2)I + 2\bar{a}\bar{a}^\top \end{pmatrix} \begin{pmatrix} x_1 \\ \bar{x} \end{pmatrix}.$$

(iii) For the EJA of m-dimensional symmetric matrices, one has that $L_A(X) = A \circ X = \frac{1}{2}(AX + XA)$ and $\mathcal{Q}_A(X) = AXA$.

By Proposition III.1.5 of [7], a Jordan algebra over $\mathbb{R}$ with a unit element $e \in \mathbb{V}$ is Euclidean iff the symmetric bilinear form $\mathrm{tr}(x \circ y)$ is positive definite. Hence, we may define an inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{V}$ by

$$\langle x, y \rangle := \mathrm{tr}(x \circ y), \quad \forall x, y \in \mathbb{V}, \tag{2}$$

and its respective norm associated by

$$\|x\| := \sqrt{\langle x, x \rangle} = \sqrt{\mathrm{tr}(x^2)} = \left(\sum_{i=1}^{r} \lambda_i^2(x) \right)^{1/2}, \quad \forall x \in \mathbb{V}.$$

We end this subsection by recalling properties that we shall employ throughout this paper. Their proofs and more details can be found in [11,41,43].

Proposition 2.2: *The following results hold:*

(a) *$x \in \mathcal{K}$ iff $\langle x, y \rangle \geq 0$ for all $y \in \mathcal{K}$. Moreover, $x \in \mathrm{int}(\mathcal{K})$ iff $\langle x, y \rangle > 0$ for all $y \in \mathcal{K} \setminus \{0\}$.*
(b) *For $x, y \in \mathcal{K}$, orthogonality condition $\langle x, y \rangle = 0$ is equivalent to $x \circ y = 0$.*
(c) *$\lambda_{\min}(x) + \lambda_{\min}(y) \leq \lambda_{\min}(x + y) \leq \lambda_{\min}(x) + \lambda_{\max}(y)$, for any $x, y \in \mathbb{V}$.*
(d) *If x is invertible, then $\mathcal{Q}_x(\mathcal{K}) = \mathcal{K}$ and $\mathcal{Q}_x(\mathrm{int}(\mathcal{K})) = \mathrm{int}(\mathcal{K})$. Moreover, $\mathcal{Q}_x$ is positive definite when $x \in \mathrm{int}(\mathcal{K})$.*
(e) *$\mathcal{Q}_{x^k} = (\mathcal{Q}_x)^k$ and $\mathcal{Q}_x(e) = x^2$ for any $x \in \mathbb{V}$ invertible.*

Unless otherwise stated, in the rest of this paper, the notation $\mathbb{V} = (\mathbb{V}, \circ, \langle \cdot, \cdot \rangle)$ represents a EJA of rank r.

3. Interior proximal algorithm with variable metric

Let us denote by $\mathcal{F} = \{x \in \mathbb{R}^n : \mathbb{A}x + b \in \mathcal{K}\}$ the feasible set of SCP problem, and by $\mathcal{F}^0 = \{x \in \mathbb{R}^n : \mathbb{A}x + b \in \mathrm{int}.\mathcal{K})\}$ its interior. Recall that $\mathbb{A}x = \sum_{k=1}^n x_k a_k$, where $a_k \in \mathbb{V}$.

Example 3.1: Some examples of the feasible set are:

(i) $Ax + b \in \mathbb{R}^m_+$, with $A = [a_1, \ldots, a_n] \in \mathbb{R}^{m \times n}$, for $\mathbb{V} = \mathbb{R}^m$ and $\mathcal{K} = \mathbb{R}^m_+$.
(ii) $Ax + b \in \mathcal{L}^m_+$, with $A = [a_1, \ldots, a_n] \in \mathbb{R}^{m \times n}$, for $\mathbb{V} = \mathbb{R}^m$ and $\mathcal{K} = \mathcal{L}^m_+$.
(iii) $w(x) = \sum_{k=1}^n x_k A_k + A_0 \in \mathcal{S}^n_+$, for $\mathbb{V} = \mathcal{S}^n$ and $\mathcal{K} = \mathcal{S}^n_+$.

From now on we suppose that the following assumptions hold true:

(A1) $f_* > -\infty$.
(A2) $\mathcal{F} \subset \mathrm{dom} f = \{x : f(x) < +\infty\}$ and $\mathcal{F}^0 \neq \emptyset$.
(A3) $\mathcal{X}^*$ (solution set of SCP problem) is nonempty and bounded.

3.1. Algorithm IPAVM

We suppose that $\mathbb{A}$ is injective. Set $\langle \cdot, \cdot \rangle_{\mathcal{Q}_{w(u)}} := \langle \mathbb{A}^\top \mathcal{Q}^{-1}_{w(u)} \mathbb{A} \cdot, \cdot \rangle$, for $u \in \mathbb{R}^n$ such that $w(u) = \mathbb{A}u + b$ be invertible, and let us define the following induced norms

$$\|v\|^2_{\mathcal{Q}_{w(u)}} := \langle v, v \rangle_{\mathcal{Q}_{w(u)}} = \langle \mathcal{Q}^{-1}_{w(u)} \mathbb{A}v, \mathbb{A}v \rangle, \tag{3}$$

and

$$\|v\|^{*2}_{\mathcal{Q}_{w(u)}} := \langle (\mathbb{A}^\top \mathcal{Q}^{-1}_{w(u)} \mathbb{A})^{-1} v, v \rangle, \tag{4}$$

for all $v \in \mathbb{R}^n$.

The interior proximal algorithm with variable metric (IPAVM) for solving the SCP problem is defined as follows:

Algorithm IPAVM: For each $k = 1, 2, \ldots$, take $\varepsilon_k > 0$, $\delta_k > 0$ and $\eta_k > 0$ with $\sum_{k=0}^{\infty} \varepsilon_k < \infty$, $\sum_{k=0}^{\infty} \delta_k < \infty$, $\sum_{k=0}^{\infty} \eta_k < \infty$ and $\{\varepsilon_k\}$ nonincreasing.

Step 0: Start with some initial point $x^0 \in \mathcal{F}^0$ and $g^0 \in \partial f(x^0)$. Set $k = 0$

Step 1: Given $x^k \in \mathcal{F}^0$, $g^k \in \partial_{\varepsilon_k} f(x^k)$, and suitable parameter $\gamma_k > 0$, find (x^{k+1}, g^{k+1}) such that

$$g^{k+1} \in \partial_{\varepsilon_{k+1}} f\left(x^{k+1}\right), \tag{5}$$

$$g^{k+1} + \gamma_k \mathbb{A}^\top \mathcal{Q}^{-1}_{w(x^k)} \mathbb{A}\left(x^{k+1} - x^k\right) = \zeta^{k+1}, \tag{6}$$

where the associated error ζ^{k+1} satisfies the following conditions:

$$\|\zeta^{k+1}\| \le \delta_k, \quad \|\zeta^{k+1}\| \max\left(\|x^{k+1}\|, \|x^k\|\right) \le \eta_k. \tag{7}$$

Step 2: If x^{k+1} satisfies a prescribed stopping rule, then stop.

Step 3: Replace k by $k+1$ and go to step 1.

Remark 1: Set $F_k(x) := f(x) + \frac{\gamma_k}{2}\|x - x^k\|^2_{\mathcal{Q}_{w(x^k)}}$. Since f is a closed proper convex function, it directly follows that F_k has bounded sublevel sets. Therefore, the optimal set of $\inf\{F_k(x)\}$ is nonempty and compact and thus there exists x^{k+1} such that (5)–(7) hold with $\varepsilon_{k+1} = \zeta^{k+1} = 0$.

3.2. Strictly feasible iterates and convergence

The following result shows that the iterations generated by the algorithm IPAVM belong to the interior of the feasible set $\mathcal{F}$, provided that γ_k is sufficiently large, which justifies the terminology: *interior proximal algorithm with variable metric* (IPAVM). In order to avoid any misleading interpretation, the smallest and the largest eigenvalues of the self-adjoint linear operator $\mathcal{Q}_z$ are denoted by bold symbols $\boldsymbol{\lambda}_{\min}(\mathcal{Q}_z)$ and $\boldsymbol{\lambda}_{\max}(\mathcal{Q}_z)$, respectively.

Proposition 3.1: *Suppose that $x^k \in \mathcal{F}^0$ for some $k \ge 0$ and the parameter γ_k satisfies $\gamma_k > \bar{\gamma}_k$, where*

$$\bar{\gamma}_k = \|\mathbb{A}^{-1}\| \boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})^{1/2}[\|g^k\| + \delta_k]. \tag{8}$$

Then, the iterate x^{k+1} generated by Step 1 of IPAVM belongs to $\mathcal{F}^0$.

Proof: Assume that $x^k \in \mathcal{F}^0$. By remark 1, there exists x^{k+1} satisfying the conditions of Step 1. From the definition of ε-subdifferential, it follows that $\langle g^{k+1} - g^k, x^{k+1} - x^k \rangle \ge \varepsilon_k - \varepsilon_{k+1}$, which together with (6) yields to

$$\gamma_k \langle \mathbb{A}^\top \mathcal{Q}^{-1}_{w(x^k)} \mathbb{A}(x^{k+1} - x^k), x^{k+1} - x^k \rangle \le \langle g^k, x^k - x^{k+1} \rangle + \langle \zeta^{k+1}, x^{k+1} - x^k \rangle + \varepsilon_{k+1} - \varepsilon_k.$$

From the Cauchy–Schwarz inequality, (7) and the fact that $\{\varepsilon_k\}$ is nonincreasing, it follows that

$$\begin{aligned}\gamma_k \left\langle \mathcal{Q}^{-1}_{w(x^k)} \mathbb{A}(x^{k+1} - x^k), \mathbb{A}(x^{k+1} - x^k) \right\rangle &\le \|g^k\| \|x^k - x^{k+1}\| + \|\zeta^{k+1}\| \|x^{k+1} - x^k\| \\ &\le [\|g^k\| + \delta_k] \|x^{k+1} - x^k\|. \end{aligned} \tag{9}$$

Since $\mathcal{Q}^{-1/2}_{w(x^k)}$ is positive definite (cf. Proposition 2.2(d)), one has that

$$\|\mathcal{Q}^{-1/2}_{w(x^k)}\mathbb{A}\left(x^{k+1}-x^k\right)\| \geq \boldsymbol{\lambda}_{\min}\left(\mathcal{Q}^{-1/2}_{w(x^k)}\right)\|\mathbb{A}\left(x^{k+1}-x^k\right)\| = \boldsymbol{\lambda}_{\min}\left(\mathcal{Q}^{-1}_{w(x^k)}\right)^{1/2}\|\mathbb{A}\left(x^{k+1}-x^k\right)\|.$$

Then,

$$\left\langle \mathcal{Q}^{-1}_{w(x^k)}\mathbb{A}(x^{k+1}-x^k), \mathbb{A}(x^{k+1}-x^k)\right\rangle \geq \boldsymbol{\lambda}_{\min}\left(\mathcal{Q}^{-1}_{w(x^k)}\right)^{1/2}\|\mathbb{A}(x^{k+1}-x^k)\| \cdot \\ \|\mathcal{Q}^{-1/2}_{w(x^k)}\mathbb{A}(x^{k+1}-x^k)\|.$$

By using this inequality in (9), it follows that

$$\gamma_k\|\mathbb{A}(x^{k+1}-x^k)\|\|\mathcal{Q}^{-1/2}_{w(x^k)}(w(x^{k+1})-w(x^k))\| \leq \boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})^{1/2}[\|g^k\|+\delta_k]\|x^{k+1}-x^k\|. \quad (10)$$

Since $\mathbb{A}$ is injective, we get $\|\mathbb{A}(x^{k+1}-x^k)\| \geq \frac{1}{\|\mathbb{A}^{-1}\|_s}\|x^{k+1}-x^k\|$, where $\|\cdot\|_s$ denotes the norm in the sense of operators. By using this inequality in (32), we obtain that

$$\|\mathcal{Q}^{-1/2}_{w(x^k)}(w(x^{k+1})-w(x^k))\| \leq \frac{1}{\gamma_k}\|\mathbb{A}^{-1}\|_s\boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})^{1/2}[\|g^k\|+\delta_k] < 1,$$

where we have used (8) in the last inequality. On the other hand, it holds from Proposition 2.2(e), that $\mathcal{Q}^{-1/2}_{w(x^k)}w(x^k) = e$, which yields $\|\mathcal{Q}^{-1/2}_{w(x^k)}(w(x^{k+1})-w(x^k))\| = \|\mathcal{Q}^{-1/2}_{w(x^k)}(x^{k+1})-e\|$, and from definition of norm on $\mathbb{V}$ it follows that

$$\|\mathcal{Q}^{-1/2}_{w(x^k)}(w(x^{k+1})-w(x^k))\| \geq |\lambda_i(\mathcal{Q}^{-1/2}_{w(x^k)}(x^{k+1})-e)|, \quad \forall i=1,\ldots,r.$$

Hence, $|\lambda_i(\mathcal{Q}^{-1/2}_{w(x^k)}(x^{k+1})-e)| \leq 1, \forall i=1,\ldots,r$, which implies in particular that

$$-1 < \lambda_{\min}(\mathcal{Q}^{-1/2}_{w(x^k)}(x^{k+1})-e) < 1.$$

By using Proposition 2.2 (c), in both inequalities, we get

$$0 = \lambda_{\min}(e)-1 < \lambda_{\min}(\mathcal{Q}^{-1/2}_{w(x^k)}(x^{k+1})) < 1+\lambda_{\max}(e) = 2.$$

This implies that $\mathcal{Q}^{-1/2}_{w(x^k)}w(x^{k+1}) \in \operatorname{int}(\mathcal{K})$, that is, $w(x^{k+1}) \in \mathcal{Q}^{1/2}_{w(x^k)}(\operatorname{int}(\mathcal{K}))$. Therefore, by Proposition 2.2(d), it follows that $x^{k+1} \in \mathcal{F}^0$. □

Let us illustrate the condition on the regularization parameter (8) in some known examples of Euclidean Jordan algebras (see Example 2.1).

Example 3.2:

(i) Let $\mathbb{V} = \mathbb{R}^m$, $\mathcal{K} = \mathbb{R}^m_+$ and $z \in \mathcal{K}$. Then, $\boldsymbol{\lambda}_{\max}(\mathcal{Q}_z)^{1/2} = \max(z)$, and

$$\bar{\gamma}_k = \max(w(x^k))[\|g^k\|+\delta_k]/\sigma_{\min}(A).$$

(ii) Let $\mathbb{V} = \mathbb{R}^m$, $\mathcal{K} = \mathcal{L}^{m_1}_+ \times \cdots \times \mathcal{L}^{m_J}_+$ and $z = (z^1,\ldots,z^J) \in \mathcal{K}$, where $m = \sum_{i=1}^J m_i$. Then, $\boldsymbol{\lambda}_{\max}(\mathcal{Q}_z)^{1/2} = \max_{i=1,\ldots,J}\{\lambda_{\max}(z^i)\}$, and

$$\bar{\gamma}_k = \max_{i=1,\ldots,J}\{\lambda_{\max}(w_i(x^k))\}[\|g^k\|+\delta_k]/\sigma_{\min}(A).$$

(iii) Let $\mathbb{V} = \mathcal{S}^n$, $\mathcal{K} = \mathcal{S}^n_+$ and $Z \in \mathcal{S}^n_+$. Then,

$$\boldsymbol{\lambda}_{\max}(\mathcal{Q}_Z)^{1/2} = \max_{U \neq 0} \frac{\|ZU\|}{\|U\|} \quad \text{and} \quad \bar{\gamma}_k = \max_{U\neq 0} \frac{\|w(x^k)U\|}{\|U\|}[\|g^k\| + \delta_k]/\sigma_{\min}(\mathbf{A}),$$

where $\mathbf{A} = [\mathrm{vec}(A_1), \ldots, \mathrm{vec}(A_n)]$, with $\mathrm{vec}(A)$ denoting the vector obtained by writing the columns of A one after another.

The following result is obtained in a way similar to [1, Proposition 3.5 and Lemma 4.2]. We include its proof for the sake of completeness.

Proposition 3.2: *Let $\{x^k\} \subset \mathcal{F}^0$ be a sequence generated by the algorithm IPAVM under $\gamma_k > \bar{\gamma}_k$ with $\bar{\gamma}_k$ given by (8) then the following hold:*

(i) *$\{f(x^k)\}$ converges and $\sum_{k=0}^{\infty} \gamma_k \|x^{k+1} - x^k\|^2_{\mathcal{Q}_{w(x^k)}} < \infty$.*

(ii) *The sequence $\{x^k\}$ is bounded.*

(iii) *If $\gamma_k \geq \bar{\gamma}_k + \beta$ for some $\beta > 0$, then $\sum_{k=0}^{\infty} \|x^{k+1} - x^k\|^2 < \infty$, and in particular, $\lim_{k\to+\infty} \|x^{k+1} - x^k\| = 0$.*

Proof:

(i) From (6) and since $g^{k+1} \in \partial_{\varepsilon_{k+1}} f(x^{k+1})$ we have $f(x^k) + \langle \zeta^{k+1}, x^{k+1} - x^k\rangle + \varepsilon_{k+1} \geq f(x^{k+1}) + \gamma_k \|x^{k+1} - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \geq f(x^{k+1})$. By (7), and using $\langle \zeta^{k+1}, x^{k+1} - x^k\rangle \leq \|\zeta^{k+1}\|(\|x^k\| + \|x^{k+1}\|) \leq 2\|\zeta^{k+1}\| \max(\|x^{k+1}\|, \|x^k\|)$, we obtain

$$f(x^{k+1}) + \gamma_k \|x^{k+1} - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \leq f(x^k) + 2\eta_k + \varepsilon_{k+1}. \tag{11}$$

Thus $0 \leq f(x^{k+1}) - f_* \leq f(x^k) - f_* + 2\eta_k + \varepsilon_{k+1}$. Hence, using [1, Lemma 3.4(i)] we deduce that the sequence $\{f(x^k)\}$ converges. On the other hand, from (11) we get $\sum_{k=0}^{N} \gamma_k \|x^{k+1} - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \leq f(x^0) - f(x^{N+1}) + \sum_{k=0}^{N} (2\eta_k + \varepsilon_{k+1}) \leq f(x^0) - f_* + \sum_{k=1}^{N+1} (2\eta_k + \varepsilon_{k+1})$. Letting $N \to +\infty$, we obtain the result.

(ii) Summing (11) over $k = 0, \ldots, l$, one has $f(x^{l+1}) - f(x^0) \leq \sum_{k=0}^{l} (2\eta_k + \varepsilon_{k+1})$. Since $\sum_{k=0}^{\infty} \eta_k$ and $\sum_{k=0}^{\infty} \varepsilon_k$ exist, it follows that for some $\bar{\eta} \geq 0$ we have $f(x^{l+1}) \leq f(x^0) + \bar{\eta} < \infty$, for all $l \geq 0$. On the other hand, from assumption (A3), we deduce that f is level bounded over $\mathcal{F}$. Thus, one has that $\{x^k\}$ is a bounded sequence.

(iii) Since $\mathcal{Q}^{-1}_{w(x^k)}$ is positive definite (cf. Proposition 2.2(d)) and $\mathbb{A}$ is injective, one has

$$\|x^{k+1} - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \geq \boldsymbol{\lambda}_{\min}(\mathcal{Q}^{-1}_{w(x^k)}) \|\mathbb{A}(x^{k+1} - x^k)\|^2 \geq \frac{1}{\|\mathbb{A}^{-1}\|_s^2 \boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})} \|x^{k+1} - x^k\|^2.$$

Now, by the boundedness of the sequence $\{x^k\}$, there exists $\tilde{\eta} > 0$ such that $\boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)}) < \tilde{\eta}$. Taking, $\tau = \frac{\beta}{\tilde{\eta}\|\mathbb{A}^{-1}\|_s^2}$, we obtain that $\sum_{k=0}^{\infty} \gamma_k \|x^{k+1} - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \geq \tau \sum_{k=0}^{\infty} \|x^{k+1} - x^k\|^2$, and the result follows from Part(i). □

Remark 2: When the function f is defined everywhere, it follows from the above proposition that $\{g^k\}$ is bounded. Moreover, that $\{\gamma_k\}$ can be chosen to be bounded.

By using the above Proposition, we can establish some partial results about of the convergence of our algorithm.

Proposition 3.3: *Suppose that f is defined in all $\mathbb{R}^n$. Let $\{x^k\} \subset \mathcal{F}^0$ be sequence generated by the algorithm IPAVM, then*

(i) *If a cluster point $\tilde{x}$ of the sequence $\{x^k\}$ belongs to $\mathcal{F}^0$, then $\tilde{x}$ is optimal for SCP.*

(ii) *The dual sequence* $\{s^{k+1}\}$ *defined by*

$$s^{k+1} := \gamma_k \mathcal{Q}^{-1}_{w(x^k)}(w(x^k) - w(x^{k+1})) \tag{12}$$

is bounded and satisfies

$$\lim_{k\to+\infty} \mathcal{Q}^{1/2}_{w(x^k)} s^{k+1} = 0. \tag{13}$$

(iii) *Any cluster point* $(\tilde{x}, \tilde{s}, \tilde{g})$ *of* $\{(x^k, s^k, g^k)\}$ *satisfies*

$$\tilde{g} = \mathbb{A}^\top \tilde{s}, \quad w(\tilde{x}) \in \mathcal{K}, \quad \langle w(\tilde{x}), \tilde{s}\rangle = 0. \tag{14}$$

Proof: Parts (i) and (ii) are followed in the same way to [1, Proposition 5.3].

(iii) By construction of sequence $\{x^k\}$, any cluster point $\tilde{x} \in \mathbb{R}^n$ satisfies $w(\tilde{x}) \in \mathcal{K}$. From (6) and (12) it follows that $\mathbb{A}^\top \tilde{s} = \tilde{g}$, with $\tilde{s}$ a limit point of the dual sequence $\{s^{k+1}\}$, which exists by Part (ii). Moreover, from (13) we get $\mathcal{Q}^{1/2}_{w(\tilde{x})}\tilde{s} = 0$. Then,

$$0 = \langle \mathcal{Q}^{1/2}_{w(\tilde{x})}\tilde{s}, e\rangle = \langle \tilde{s}, \mathcal{Q}^{1/2}_{w(\tilde{x})} e\rangle = \langle \tilde{s}, w(\tilde{x})\rangle,$$

where the second equality it follows from the fact that Q_x is self-adjoint, and the third one from Proposition 2.2(e). □

Remark 3: Note that in order to verify that any cluster point $(\tilde{x}, \tilde{s}, \tilde{g})$ of the sequence $\{(x^k, s^k, g^k)\}$ satisfies the Karush–Kuhn–Tucker (KKT) conditions of the problem SCP, it only remains to prove that $\tilde{s} \in \mathcal{K}$. First, we suppose that $w(\tilde{x}) \in \operatorname{int}(\mathcal{K})$, then $\mathcal{Q}^{1/2}_{w(\tilde{x})}$ is positive definite (cf. Proposition 2.2(d)) and hence $\mathcal{Q}^{1/2}_{w(\tilde{x})}\tilde{s} = 0$ implies that $\tilde{s} = 0$. Second, we consider the case when $w(\tilde{x}) \in \operatorname{bd}(\mathcal{K}) \setminus \{0\}$, with $\operatorname{bd}(\mathcal{K})$ denoting the boundary of $\mathcal{K}$. We argue by contradiction, that is, we suppose that $\tilde{s} \in -\operatorname{int}(\mathcal{K})$. By Proposition 2.2(a) we get $\langle w(\tilde{x}), \tilde{s}\rangle < 0$, which is a contradiction. Hence, $\tilde{s} \notin -\operatorname{int}(\mathcal{K})$. This implies that $\lambda_{\max}(\tilde{s}) \geq 0$.

3.3. Particular cases of convergence

In this section, we present some particular cases where we have been able to establish that any cluster point of the sequence generated by our algorithm IPAVM, satisfies the KKT conditions.

In the first case, we assume that the objective function is linear. Then, by using the similar arguments that in [1, Proposition 5.5], given for Euclidean Jordan algebra of quadratic forms, we obtain the following result.

Proposition 3.4: *Under the assumptions and notations of Proposition 3.3, if in addition f is supposed to be $f(x) = \langle c, x\rangle$, then any limit point of $\{x^k\}$ satisfies the KKT conditions.*

Now, in the second case, we consider Euclidean Jordan algebra of m-dimensional vectors, and $w(x) = Ax$ with $A \in \mathbb{R}^{m\times n}$. Note that in this case, the relation (6) is written as

$$g^{k+1} = A^\top s^{k+1} + \zeta^{k+1}, \quad s^{k+1} = \gamma_k[\operatorname{Diag}(w(x^k)^2)]^{-1}(w(x^k) - w(x^{k+1})). \tag{15}$$

Then, we by using the ideas from [44, Lemma 1] we get the following result.

Proposition 3.5: *Any limit point $(\tilde{x}, \tilde{g}, \tilde{s})$ of the sequence $\{(x^k, g^k, s^k)\}$ satisfies the KKT conditions*

$$\tilde{g} = A^\top \tilde{s},\ A\tilde{x} \in \mathbb{R}^m_+,\ \tilde{s} \in \mathbb{R}^m_+, \quad (A\tilde{x})_i \tilde{s}_i = 0,\ \text{for} i = 1, \ldots, m,$$

where $\tilde{g} \in \partial f(\tilde{x})$.

Proof: Let $(\tilde{x}, \tilde{g}, \tilde{s})$ be a limit point of $\{(x^k, g^k, s^k)\}$. In order to verify that $(\tilde{x}, \tilde{g}, \tilde{s})$ satisfies KKT condition, it only remains to prove that $\tilde{s} \in \mathbb{R}^m_+$ (cf. Proposition 3.3). We argue by contradiction. Suppose that there exists i such that $\tilde{s}_i < 0$. We define the following sets

$$I = \{i : \tilde{s}_i < 0\}, \quad J = \{1, \dots, m\} \setminus I, \quad V = \{x \in \mathbb{R}^n : Ax \geq 0, (Ax)_i = 0, i \in I\}.$$

Clearly $\tilde{x} \in V$ (cf. (14)). We claim that $\tilde{x}$ is an optimal solution of the convex problem

$$\min\{f(x) : x \in V\}. \tag{16}$$

Indeed, the KKT conditions for problem (16) are given by

$$g = \sum_{j \in J} (A^j)^\top s_j, \quad (Ax)_j \geq 0, \quad s_j \geq 0, \quad (Ax)_j s_j = 0, \quad \forall j \in J \text{ with } g \in \partial f(x), \tag{17}$$

where $s \in \mathbb{R}^{|J|}$ and A^j denotes the j-row of A. These conditions are obviously satisfied by $\tilde{x}$ with $\tilde{g} \in \partial f(\tilde{x})$ and $(\tilde{s}_j)_{j \in J}$. Let

$$B = \{(x, s) \in \mathbb{R}^n \times \mathbb{R}^m : Ax \geq 0, s_i < 0, i \in I\}.$$

Clearly $(\tilde{x}, \tilde{s}) \in B$. Since $\{(x^k, g^k, s^k)\}$ is bounded, we can take a subsequence $\{(x^{l_k}, g^{l_k}, s^{l_k})\}$ such that $x^{l_k} \to \tilde{x}$, $g^{l_k} \to \tilde{g}$ and $s^{l_k} \to \tilde{s}$ as $k \to \infty$, and $(x^{l_k}, s^{l_k}) \in B$ for all k. Let us define

$$p_k := \max\{q < l_k : (x^{q+1}, s^{q+1}) \notin B\}, \tag{18}$$

and set $p_k = 0$ if $(x^{q+1}, s^{q+1}) \in B$ for all $q < l_k$. We immediately get $(x^{q+1}, s^{q+1}) \in B$ for all $q = p_k + 1, \dots, l_k$, that is, $Ax^{q+1} \geq 0$, $s_i^{q+1} < 0$ for all $q = p_k + 1, \dots, l_k$ and for all $i \in I$. Hence, from (15) we obtain

$$(Ax^{q+1})_i \geq (Ax^q)_i, \quad \forall q = p_k + 1, \dots, l_k \text{ and } \quad \forall i \in I.$$

Then, by applying the above inequality iteratively, we get

$$(Ax^{l_k})_i \geq (Ax^{p_k+1})_i, \quad \forall k \text{ and } \forall i \in I.$$

Note that if there exists p such that $p_k = p$ for all k large enough, then

$$0 = (A\tilde{x})_i = \lim_{k \to \infty} (Ax^{l_k})_i \geq \lim_{k \to \infty} (Ax^{p_k+1})_i = (Ax^{p+1})_i > 0, \quad \forall i \in I$$

(where the last inequality it follows from Proposition 3.1), obtaining a contradiction. Thus, it follows that $\{p_k\} \to \infty$ and hence $\{x^{p_k}\}$ is a subsequence of $\{x^k\}$.

As before, we get that

$$0 = (A\tilde{x})_i = \lim_{k \to \infty} (Ax^{l_k})_i \geq \lim_{k \to \infty} (Ax^{p_k+1})_i \geq 0, \quad \forall i \in I,$$

whence $\lim_{k \to \infty} (Ax^{p_k+1})_i = 0, \forall i \in I$. By using Proposition 3.2(iii) one obtain that $\lim_{k \to \infty} (Ax^{p_k})_i = 0, \quad \forall i \in I$. Moreover, since $\{(x^{p_k}, s^{p_k})\}$ is bounded we can suppose that $(x^{p_k}, s^{p_k}) \to (\hat{x}, \hat{s})$ as $k \to \infty$, for some $(\hat{x}, \hat{s}) \notin B$ (note that $(x^{p_k}, s^{p_k}) \notin B$ for all k and that set B is open in $\mathbb{R}^n \times \mathbb{R}^m_+$). Then, $(A\hat{x})_i = \lim_{k \to \infty} (Ax^{p_k})_i = 0, \forall i \in I$, and there exists $r \in I$ such that $\hat{s}_r \geq 0$. We claim that $\hat{x}$ is also a solution of problem (16). Indeed, it is clear that $\hat{x} \in V$ and $f(\hat{x}) = f(\tilde{x}) = f^*$ (cf.

Proposition 3.2(i)). Since $\tilde{x}$ minimizes f on V, we conclude that $\hat{x}$ also minimizes f on V, that is, $(\hat{x}, \hat{s})$ satisfies (17). Also, it satisfies $(A\hat{x})_r \geq 0$, $\hat{s}_r \geq 0$ and $\hat{s}_r(A\hat{x})_r = 0$. Hence, $\hat{x}$ minimizes f on $W = \{x \in \mathbb{R}^n : Ax \geq 0, (Ax)_i = 0, i \in I \setminus \{r\}\}$. Since $\tilde{x} \in V \subset W$ and $f(\hat{x}) = f(\tilde{x})$, it follows that $\tilde{x}$ also minimizes f on W. The KKT conditions of this problem implies in particular that $\tilde{s}_r \geq 0$. But $\tilde{s}_r < 0$ when $r \in I$, obtaining a contradiction. Therefore, $\tilde{s} \geq 0$. □

4. Interior proximal bundle algorithm with variable metric

In this section, we describe the feasible bundle method to solve the nondifferentiable convex problem under symmetric cone constraints.

Recall that for an arbitrary point $x^0 \in \mathcal{F}^0$, the exact version of algorithm IPAVM generates recursively the auxiliary problem

$$x^{k+1} = \operatorname*{argmin}_{x\in\mathbb{R}^n} \left\{ f(x) + \frac{\gamma_k}{2} \|x - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \right\}, \tag{19}$$

where $\gamma_k > \bar{\gamma}_k$, with $\bar{\gamma}_k$ defined in (8) and $\|\cdot\|_{\mathcal{Q}_{w(x^k)}}$ defined in (3). This problem can be difficult to solve, whenever f is nondifferentiable. In this case, we can work with the *bundle methods*. At iteration k we have at our disposal the sequence $x^0, x^1, \ldots, x^k$. Let $J^\ell \subseteq \{0, 1, \ldots, \ell\} \subset \mathbb{N}$ be a finite index set and y^j $(j \in J^\ell)$ be arbitrary points (we can assume that $y^j = x^j$). With the function values $f(y^j)$ and arbitrary $g^j \in \partial f(y^j)$ we define for each $j \in J^\ell$ the *linearization* of f at y^j as

$$\breve{f}_j(x) := f(y^j) + \langle g^j, x - -y^j\rangle$$

and the (nonnegative, by convexity) *linearization error* at x^k by

$$e_j := f(x^k) - \breve{f}_j(x^k) = f(x^k) - (f(y^j) + \langle g^j, x^k - y^j\rangle), \quad \forall j \in J^\ell.$$

The set $\mathcal{B}_\ell = \{(y^j, f(y^j), g^j) : j \in J^\ell\}$ is called *the bundle*. With the bundle of past information at each iteration a *cutting-planes model* of the objective function f is defined by

$$\varphi_\ell(y) = \max_{j\in J^\ell} \breve{f}_j(y) = \max_{j\in J^\ell} \{f(y^j) + \langle g^j, y - y^j\rangle\}, \quad \forall y \in \mathbb{R}^n. \tag{20}$$

This function is a piecewise linear approximation from below of the convex function f, that is, $\varphi_\ell \leq f$ and is used to generate a point $y^{\ell+1}$. It is worthwhile to note that

$$\varphi_\ell(y^j) = f(y^j), \quad \text{for all } j \in J^\ell. \tag{21}$$

Then, replacing f by φ_ℓ in (19), we obtain the following quadratic minimization problem

$$\min_{y\in\mathbb{R}^n} \left\{ \varphi_\ell(y) + \frac{\gamma_k}{2} \|y - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \right\}, \tag{22}$$

which gives a unique optimal solution $y^{\ell+1}$. Note that the bundle $\mathcal{B}_\ell$ is generated using a prox-center or stability center x^k with $j \leq \ell$.

With an additional variable $r \in \mathbb{R}$ the last problem can equivalently be written as

$$\begin{array}{c} \min\limits_{(r,y)\in\mathbb{R}^{n+1}} \left\{ r + \frac{\gamma_k}{2} \|y - x^k\|^2_{\mathcal{Q}_{w(x^k)}} \right\} \\ \text{s:t: } f(y^j) + \langle g^j, y - y^j\rangle \leq r, \ \forall j \in J^\ell. \end{array} \tag{23}$$

Note that the dual of (22) is a convex quadratic problem of the following form

$$\begin{aligned} \min_{\alpha\in\mathbb{R}^{|J^\ell|}} & \left\{ \frac{1}{2}\left\| \sum_{j\in J^\ell} \alpha_j g^j \right\|^{*2}_{\mathcal{Q}_{w(x^k)}} + \gamma_k \sum_{j\in J^\ell} \alpha_j e_j \right\} \\ \text{s:t:} & \sum_{j\in J^\ell} \alpha_j = 1,\ \alpha_j \geq 0, \quad \forall j \in J^\ell, \end{aligned} \tag{24}$$

where $\|\cdot\|^*_{\mathcal{Q}_{w(x^k)}}$ is defined in (4).

As consequence of [29, Lemma 9.8] we have that, if the solution $\{\alpha_j^k : j \in J^\ell\}$ of (24) is calculated, the solution $y^{\ell+1}$ of (22) can be obtained by

$$y^{\ell+1} = x^k - \gamma_k^{-1}(\mathbb{A}^\top \mathcal{Q}^{-1}_{w(x^k)}\mathbb{A})^{-1}\tilde{g}^\ell, \quad \text{where } \tilde{g}^\ell = \sum_{j\in J^\ell} \alpha_j^k g^j \in \partial\varphi_\ell(y^{\ell+1}). \tag{25}$$

Moreover, it follows from [29, Lemma 9.8(iii)] that $\tilde{g}^\ell \in \partial_{\varepsilon_\ell} f(x^k)$ where

$$\varepsilon_\ell = f(x^k) - \varphi_\ell(y^{\ell+1}) - \frac{1}{\gamma_k}\|\tilde{g}^\ell\|^{*2}_{\mathcal{Q}_{w(x^k)}} \geq 0,$$

which is called *aggregate error*.

We are now ready to provide the proximal bundle algorithm with variable metric for solving the problem with symmetric cone constraints.

Algorithm IPBAVM: For each $k = 1, 2, \ldots$, generate the sequence $\{x^k\} \subset \mathcal{F}^0$ as follows:.

Step 0: Choose an $m \in (0, 1)$. Select starting point $x^0 \in \mathcal{F}^0$, $g^0 \in \partial f(x^0)$ and suitable parameter $\gamma_0 > 0$. Set $y^0 = x^0$, $J^0 = \{0\}$, $e_0 = 0$, and set the counters $\ell = 0$, $k = 0$.

Step 1: Find multiplier α_j^k $(j \in J^\ell)$ that it solves the dual problem (24). Set $\hat{J}^\ell = \{j \in J^\ell : \alpha_j^k \neq 0\}$. Calculate

$$\tilde{g}^\ell = \sum_{j\in\hat{J}^\ell} \alpha_j^k g^j \quad \text{and} \quad \varepsilon_\ell = \sum_{j\in\hat{J}^\ell} \alpha_j^k e_j,$$

and compute

$$\delta_\ell = \varepsilon_\ell + \frac{1}{\gamma_k}\|\tilde{g}^\ell\|^{*2}_{\mathcal{Q}_{w(x^k)}} \geq 0. \tag{26}$$

Step 2: Set $y^{\ell+1} = x^k - \gamma_k^{-1}(\mathbb{A}^\top \mathcal{Q}^{-1}_{w(x^k)}\mathbb{A})^{-1}\tilde{g}^\ell$ with $\gamma_k > \bar{\gamma}_k$, and $\bar{\gamma}_k$ defined in (8).

IF (Descent test)

$$f(y^{\ell+1}) \leq f(x^k) - m\delta_\ell = f(x^k) - m(f(x^k) - \varphi_\ell(y^{\ell+1})),$$

THEN (Serious step)

set $x^{k+1} = y^{\ell+1}$. If x^{k+1} satisfies a prescribed stopping rule, then stop.

Else, choose $g^{\ell+1} \in \partial f(x^{k+1})$.

Linearization error update

$$\begin{aligned} e_j &= e_j + f(x^{k+1}) - f(x^k) - \langle g^j, x^{k+1} - x^k\rangle, \quad \forall j \in J^\ell, \\ e_{\ell+1} &= 0. \end{aligned}$$

Update $\gamma_{k+1} > 0$. Replace k by $k + 1$.

ELSE (Null step)

choose $g^{\ell+1} \in \partial f(y^{\ell+1})$.
Linearization error update

$$e_j = e_j, \quad \forall j \in J^{\ell},$$
$$e_{\ell+1} = f(x^k) - f(y^{\ell+1}) + \langle g^{\ell+1}, y^{\ell+1} - x^k \rangle,$$

Step 3: $J^{\ell+1} := \hat{J}^{\ell} \cup \{\ell+1\}$, increase ℓ by 1 and go to step 1.

Remark 4: Note that φ_ℓ given by (20) can be replaced with any model function satisfying the following properties:

$$\varphi_\ell \le f, \text{ for } \ell = 0, 1, \ldots, \tag{27}$$
$$l^{\ell} \le \varphi_{\ell+1}, \tag{28}$$
$$f(y^\ell) + \langle g^\ell, y - y^\ell \rangle \le \varphi_\ell(y), \; \forall\, y \text{ with } g^\ell \in \partial f(y^\ell), \tag{29}$$

where

$$l^{\ell}(y) = \varphi_\ell(y^{\ell+1}) + \langle \tilde{g}^\ell, y - y^{\ell+1} \rangle, \tag{30}$$

called the *aggregate linearization* of f, with $\tilde{g}^\ell \in \partial \varphi_\ell(y^{\ell+1})$. From (25) it follows that the aggregate error can be written as $\varepsilon_\ell = f(x^k) - l^\ell(x^k)$. Some examples of φ_ℓ satisfying (27)–(29) are given in [30].

The following result shows that the bundle subiterations belong to the interior of the feasible set, and that this holds true for an arbitrary model function φ_ℓ satisfying the properties (27)–(29).

Proposition 4.1: *The subiterations $\{y^\ell\}$ generated by IPBAVM belong to $\mathcal{F}^0$.*

Proof: We assume that $x^k = y^{\ell_{k-1}}$ satisfies the serious step and belong to $\mathcal{F}^0$. Let $y^{j+1} = x^k - \gamma_k^{-1}(\mathbb{A}^\top \mathcal{Q}^{-1}_{w(x^k)}\mathbb{A})^{-1}\tilde{g}^j$, with $\tilde{g}^j \in \partial\varphi_j(y^{j+1})$, for $j = \ell_{k-1}+1, \ldots, \ell_k$. We suppose that for $j = \ell_{k-1} + 1, \ldots, \ell_k$ the serious step fails and that for $j = \ell_k + 1$ the serious step holds. We prove that $y^{j+1} \in \mathcal{F}^0$, for all $j = \ell_{k-1}+1, \ldots, \ell_k$. From the monotonicity of $\partial\varphi_j$, it follows that $\langle \tilde{g}^j - \bar{g}^k, y^{j+1} - x^k \rangle \ge 0$, where $\bar{g}^k \in \partial\varphi_j(x^k)$. Replacing the value of $\tilde{g}^j$, and then by using the Cauchy–Schwarz inequality, we obtain

$$\gamma_k \langle \mathcal{Q}^{-1}_{w(x^k)} \mathbb{A}(y^{j+1} - x^k), \mathbb{A}(y^{j+1} - x^k) \rangle \le \langle \bar{g}^k, x^k - y^{j+1} \rangle \le \|\bar{g}^k\| \|y^{j+1} - x^k\|, \tag{31}$$

Now, since $\mathcal{Q}^{-1/2}_{w(x^k)}$ is positive definite (cf. Proposition 2.2(d)), one has that

$$\|\mathcal{Q}^{-1/2}_{w(x^k)}\mathbb{A}(y^{j+1} - x^k)\| \ge \boldsymbol{\lambda}_{\min}(\mathcal{Q}^{-1/2}_{w(x^k)})\|\mathbb{A}(y^{j+1} - x^k)\| = \boldsymbol{\lambda}_{\min}(\mathcal{Q}^{-1}_{w(x^k)})^{1/2}\|\mathbb{A}(y^{j+1} - x^k)\|.$$

Then,

$$\|\mathcal{Q}^{-1/2}_{w(x^k)}\mathbb{A}(y^{j+1} - x^k)\|^2 \ge \boldsymbol{\lambda}_{\min}(\mathcal{Q}^{-1}_{w(x^k)})^{1/2}\|\mathbb{A}(y^{j+1} - x^k)\|\|\mathcal{Q}^{-1/2}_{w(x^k)}\mathbb{A}(y^{j+1} - x^k)\|.$$

By using this inequality in (31), it follows that

$$\gamma_k \|\mathbb{A}(y^{j+1} - x^k)\| \|\mathcal{Q}^{-1/2}_{w(x^k)}(w(y^{j+1}) - w(x^k))\| \le \boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})^{1/2}\|\bar{g}^k\|\|y^{j+1} - x^k\|. \tag{32}$$

Since $\mathbb{A}$ is injective, we get $\|\mathbb{A}(y^{j+1} - x^k)\| \ge \frac{1}{\|\mathbb{A}^{-1}\|_s}\|y^{j+1} - x^k\|$, where $\|\cdot\|_s$ denotes the norm in the sense of operators. By using this inequality in (32), we obtain that

$$\gamma_k \|\mathcal{Q}^{-1/2}_{w(x^k)}(w(y^{j+1}) - w(x^k))\| \le \|\mathbb{A}^{-1}\|_s \boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})^{1/2}\|\bar{g}^k\|. \tag{33}$$

On the other hand, as $\bar{g}^k \in \partial\varphi_j(x^k)$, that is, $\varphi_j(x^k) + \langle \bar{g}^k, z - x^k\rangle \le \varphi_j(z), \forall z \in \mathbb{R}^n$, (21) holds with $x^k = y^{\ell_{k-1}}$, and φ_j is minorant of f (cf. (27)), we get

$$f(x^k) + \langle \bar{g}^k, z - x^k\rangle \le f(z), \quad \forall z \in \mathbb{R}^n$$

whence $\bar{g}^k \in \partial f(x^k)$. Thus, we can take $\bar{g}^k = g^k$. Then, by using this and the fact that $\gamma_k > \bar{\gamma}_k$ we obtain from (33) that

$$\|\mathcal{Q}^{-1/2}_{w(x^k)}(w(y^{j+1}) - w(x^k))\| < 1.$$

Note that $\mathcal{Q}^{-1/2}_{w(x^k)} w(x^k) = e$ (cf. Proposition 2.2(e)), and

$$\|\mathcal{Q}^{-1/2}_{w(x^k)} w(y^{j+1}) - e\| \ge |\lambda_i(\mathcal{Q}^{-1/2}_{w(x^k)}(y^{j+1}) - e)|, \quad \forall i = 1, \dots, r.$$

Then, from the above inequalities we get that

$$-1 < \lambda_{\min}(\mathcal{Q}^{-1/2}_{w(x^k)} w(y^{j+1}) - e) < 1,$$

for all $j = \ell_{k-1} + 1, \dots, \ell_k$. By using Proposition 2.2(c), in both inequalities, we get $0 < \lambda_{\min}(\mathcal{Q}^{-1/2}_{w(x^k)} w(y^{j+1})) < 2$, for all $j = \ell_{k-1} + 1, \dots, \ell_k$. This implies that $\mathcal{Q}^{-1/2}_{w(x^k)} w(y^{j+1}) \in \operatorname{int}(\mathcal{K})$, that is, $w(y^{j+1}) \in \mathcal{Q}_{w(x^k)^{1/2}}(\operatorname{int}(\mathcal{K}))$ for all $j = \ell_{k-1} + 1, \dots, \ell_k$. Therefore, by Proposition 2.2(d) it follows that $y^{j+1} \in \mathcal{F}^0$, for all $j = \ell_{k-1} + 1, \ell_{k-1} + 2, \dots, \ell_k$. □

4.1. Convergence of IPBAVM

In this section, we give some results of convergence for the algorithm IPBAVM.

First, we prove that the bundle subiterations converge to a proximal point, for arbitrary φ_ℓ satisfying the properties (27)–(29). Without loss of generality, let us denote by $\operatorname{prox}_{\bar{\gamma}f}(\bar{x}) \in \mathcal{F}^0$ the unique solution of problem (19) with prox-center $\bar{x} \in \mathcal{F}^0$ and parameter $\bar{\gamma} > 0$. The following result is an extension of [37, Lemma 8] and [36, Theorem 3.1] to our context.

Theorem 4.2: *Suppose that f is finite. Let $\{y^{\ell+1}\}$ be the bundle subiterations generated by the problem* (22) *with prox-center $\bar{x}$, where the descent test is omitted, i.e. only null steps are made. Set $\xi_{\ell+1} = f(y^{\ell+1}) - \varphi_\ell(y^{\ell+1})$. Then the following hold:*

(i) $\tilde{g}^\ell \in \partial_{\xi_{\ell+1}} f(y^{\ell+1})$.
(ii) *The sequence $\{y^{\ell+1}\}$ is bounded.*
(iii) $\xi_{\ell+1} \ge 0, \forall \ell \ge 0$ *and* $\lim_{\ell\to+\infty} \xi_{\ell+1} = 0$.
(iv) *The sequence $\{y^{\ell+1}\}$ converges to* $\operatorname{prox}_{\bar{\gamma}f}(\bar{x})$.

Proof:

(i) From (25) and (27), one has that

$$\langle \tilde{g}^\ell, y - y^{\ell+1}\rangle \le \varphi_\ell(y) - \varphi_\ell(y^{\ell+1}) \le f(y) - \varphi_\ell(y^{\ell+1}) = f(y) - f(y^{\ell+1}) + \xi_{\ell+1}.$$

Hence, the result follows.

(ii) For all ℓ and $y \in \mathcal{F}^0$, we define the following functions

$$\hat{l}^\ell(y) = l^\ell(y) + \frac{\bar{\gamma}}{2}\|y - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}}, \qquad \hat{\varphi}_\ell(y) = \varphi_\ell(y) + \frac{\bar{\gamma}}{2}\|y - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}},$$

with l^ℓ defined in (30).

By using (27), the definition of $y^{\ell+1}$ and the fact that $l^\ell(y^{\ell+1}) = \varphi_\ell(y^{\ell+1})$, we get

$$f(\bar{x}) \geq \varphi_\ell(\bar{x}) = \hat{\varphi}_\ell(\bar{x}) \geq \hat{\varphi}_\ell(y^{\ell+1}) = \varphi_\ell(y^{\ell+1}) + \frac{\bar{\gamma}}{2}\|y^{\ell+1} - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}}$$
$$= l^\ell(y^{\ell+1}) + \frac{\bar{\gamma}}{2}\|y^{\ell+1} - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}} = \hat{l}^\ell(y^{\ell+1}). \quad (34)$$

Also, from (28) we obtain

$$\hat{\varphi}_\ell(y^{\ell+1}) \geq l^{\ell-1}(y^{\ell+1}) + \frac{\bar{\gamma}}{2}\|y^{\ell+1} - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}} = \hat{l}^{\ell-1}(y^{\ell+1}). \quad (35)$$

On the other hand,

$$\hat{l}^{\ell-1}(y) - \hat{l}^{\ell-1}(y^\ell) = \langle \tilde{g}^{\ell-1}, y - y^\ell\rangle + \frac{\bar{\gamma}}{2}(\|y - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}} - \|y^\ell - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}})$$
$$= \langle \bar{\gamma}(\mathbb{A}^\top \mathcal{Q}^{-1}_{w(\bar{x})}\mathbb{A})(\bar{x} - y^\ell), y - y^\ell\rangle + \frac{\bar{\gamma}}{2}(\|y - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}} - \|y^\ell - \bar{x}\|^2_{\mathcal{Q}_{w(\bar{x})}})$$
$$= \frac{\bar{\gamma}}{2}\|y - y^\ell\|^2_{\mathcal{Q}_{w(\bar{x})}}, \quad (36)$$

where the first equality is due to definition of l^ℓ, the second one to (25) and the third one to property of difference of norms. Now, setting $y^{\ell+1}$ for y in (36) one has that

$$\hat{l}^{\ell-1}(y^{\ell+1}) - \hat{l}^{\ell-1}(y^\ell) = \frac{\bar{\gamma}}{2}\|y^{\ell+1} - y^\ell\|^2_{\mathcal{Q}_{w(\bar{x})}}. \quad (37)$$

Then, combining (34), (35) and (37) gives

$$f(\bar{x}) \geq \hat{l}^\ell(y^{\ell+1}) \geq \hat{l}^{\ell-1}(y^\ell) + \frac{\bar{\gamma}}{2}\|y^{\ell+1} - y^\ell\|^2_{\mathcal{Q}_{w(\bar{x})}} \geq \hat{l}^{\ell-1}(y^\ell).$$

Hence, the sequence $\{\hat{l}^{\ell-1}(y^\ell)\}_{\ell\geq 1}$ is bounded and monotone, that is, it converges to some $l^* \in \mathbb{R}$. Moreover, from (27)–(28) one has $l^{\ell-1}(\bar{x}) \leq f(\bar{x})$, which is equivalent to $\hat{l}^{\ell-1}(\bar{x}) \leq f(\bar{x})$. Setting $\bar{x}$ for y in (36), using the fact that $\{\hat{l}^{\ell-1}(y^\ell)\}_{\ell\geq 1}$ converges to l^* and that $\bar{\gamma}$ is bounded (cf. Remark 2), it follows that the sequence $\{y^\ell\}$ is bounded.

(iii) By Proposition 4.1, we get that $y^\ell \in \mathcal{F}^0 \subset \mathrm{int}(\mathrm{dom} f)$. Then, the mean value theorem implies the existence of z^ℓ in the open line segment $(y^\ell, y^{\ell+1})$ with $c^\ell \in \partial f(z^\ell)$ such that

$$f(y^{\ell+1}) - f(y^\ell) = \langle c^\ell, y^{\ell+1} - y^\ell\rangle. \quad (38)$$

Since f is defined everywhere, it follows that the sequences $\{c^\ell\}$ and $\{g^\ell\}$ are bounded. Note also that from (27), (29), (38), and the Cauchy–Schwartz inequality, we obtain

$$\|c^\ell\|\|y^{\ell+1} - y^\ell\| \geq f(y^{\ell+1}) - f(y^\ell) \geq \varphi_\ell(y^{\ell+1}) - f(y^\ell) \geq \langle g^\ell, y^{\ell+1} - y^\ell\rangle \geq -\|g^\ell\|\|y^{\ell+1} - y^\ell\|.$$

On the other hand, using the fact that $\{\hat{l}^{\ell-1}(y^\ell)\}_{\ell\geq 1}$ converges and that $\bar{\gamma}$ is bounded (cf. Remark 2), we get from (37) that

$$\lim_{\ell\to+\infty} \|y^{\ell+1} - y^\ell\|_{\mathcal{Q}_{w(\bar{x})}} = 0. \quad (39)$$

By using (39) in the above inequality yields

$$\lim_{\ell\to+\infty}(f(y^{\ell+1})-f(y^{\ell}))=\lim_{\ell\to+\infty}(\varphi_{\ell}(y^{\ell+1})-f(y^{\ell}))=0.$$

In consequence, as

$$0\le\xi_{\ell+1}=f(y^{\ell+1})-\varphi_{\ell}(y^{\ell+1})=f(y^{\ell+1})-f(y^{\ell})+f(y^{\ell})-\varphi_{\ell}(y^{\ell+1}),$$

the result follows.

(iv) By Part (i) and (25) we get

$$0\in\partial_{\xi_{\ell+1}}f(y^{\ell+1})+\bar{\gamma}(\mathbb{A}^{\top}\mathcal{Q}_{w(\bar{x})}^{-1}\mathbb{A})(y^{\ell+1}-\bar{x}),$$

and from definition of $\mathrm{prox}_{\bar{\gamma}f}(\bar{x})$ one has that

$$0\in\partial f(\mathrm{prox}_{\bar{\gamma}f}(\bar{x}))+\bar{\gamma}(\mathbb{A}^{\top}\mathcal{Q}_{w(\bar{x})}^{-1}\mathbb{A})(\mathrm{prox}_{\bar{\gamma}f}(\bar{x})-\bar{x}).$$

Then, from the above inclusions it follows that

$$\begin{aligned}\xi_{\ell+1}&\ge\bar{\gamma}\langle(\mathbb{A}^{\top}\mathcal{Q}_{w(\bar{x})}^{-1}\mathbb{A})(\mathrm{prox}_{\bar{\gamma}f}(\bar{x})-y^{\ell+1}),\mathrm{prox}_{\bar{\gamma}f}(\bar{x})-y^{\ell+1}\rangle\\&\ge\bar{\gamma}\boldsymbol{\lambda}_{\min}(\mathbb{A}^{\top}\mathcal{Q}_{w(\bar{x})}^{-1}\mathbb{A})\|\mathrm{prox}_{\bar{\gamma}f}(\bar{x})-y^{\ell+1}\|^2,\end{aligned}$$

where the second inequality it follows from the injectivity of $\mathbb{A}$. Finally, letting $\ell\to+\infty$ and by using Part (iii), we get that $y^{\ell+1}\to\mathrm{prox}_{\bar{\gamma}f}(\bar{x})$. □

The next result extends [37, Theorem 4], [36, Theorem 3.2(a)] to our context.

Theorem 4.3: *Suppose that f is finite. Let $\{x^k\}$ be the sequence generated by the algorithm IPBAVM. If x^k is not optimal solution, then the bundle subiterations gives in a finite number of steps $\ell(k)$ a point $y^{\ell(k)+1}$ satisfying the descent test.*

Proof: Let $\{y^{\ell+1}\}$ be the sequence generated by the bundle subiterations with prox-center x^k. We denote by $z_k=\mathrm{prox}_{\gamma_k f}(x^k)\in\mathcal{F}^0$ and $\Phi(y)=f(y)+\frac{\gamma_k}{2}\|y-x^k\|^2_{\mathcal{Q}_{w(x^k)}}$. By Theorem 4.2 (iv), $y^{\ell+1}\to z_k$ as $\ell\to+\infty$. Then, by continuity of f at z_k it follows that $f(y^{\ell+1})\to f(z_k)$ as $\ell\to\infty$. On the other hand, suppose that x^k is not optimal solution for SCP and that the descent test fails at step k. Then, for each ℓ one has

$$f(y^{\ell+1})-f(x^k)>-m(f(x^k)-\varphi_{\ell}(y^{\ell+1}))=m(\varphi_{\ell}(y^{\ell+1})-f(y^{\ell+1})+f(y^{\ell+1})-f(x^k)),$$

thus

$$(1-m)(f(y^{\ell+1})-f(x^k))>-m\xi_{\ell+1}.$$

Letting $\ell\to+\infty$ in this inequality, and by using Theorem 4.2 (iii), we get that $f(z_k)=\liminf f(y^{\ell+1})\ge f(x^k)$. Hence, $\Phi(z_k)\ge f(x^k)=\Phi(x^k)$ and from definition of z_k it follows that $z_k=x^k$. Finally, since $\emptyset\ne\mathcal{F}^0\subset\mathrm{int}(\mathrm{dom}(f))$, we can apply [40, Theorem 23.8] so that $0\in\partial\Phi(x^k)=(\partial f(\cdot)+\gamma_k(\mathbb{A}^{\top}\mathcal{Q}_{w(x^k)}^{-1}\mathbb{A})(\cdot-x^k))(x^k)=\partial f(x^k)$, which means that x^k is an optimal solution of SCP problem, a contradiction. □

In the following result, we show that the descend steps fit the framework of Section 3.

Theorem 4.4: *Suppose that f is finite, and that infinitely many descent steps occur. Then,*

(i) $\sum_{k=0}^{\infty}\xi_{k+1}<\infty$.
(ii) $\sum_{k=0}^{\infty}\gamma_k\|x^{k+1}-x^k\|^2_{\mathcal{Q}_{w(x^k)}}<\infty$, *the sequence $\{x^k\}$ is bounded and $\{f(x^k)\}$ converges.*

(iii) *Any cluster point* $(\tilde{x}, \tilde{s}, \tilde{g})$ *of* $\{x^k, s^k, \tilde{g}^k\}$ *satisfies*

$$\tilde{g} = \mathbb{A}^\top \tilde{s}, \quad w(\tilde{x}) \in \mathcal{K}, \quad \langle w(\tilde{x}), \tilde{s}\rangle = 0, \tag{40}$$

where $s^k = \gamma_k \mathcal{Q}^{-1}_{w(x^k)}(w(x^k) - w(x^{k+1}))$.

Proof: (i) At iteration k, suppose that x^k was not optimal solution. By Theorem 4.3, we compute a point $y^{\ell(k)+1}$ such that it satisfies the descent test, so we can take $x^{k+1} = y^{\ell(k)+1}$. Then,

$$f(x^{k+1}) - f(x^k) \le -m(f(x^k) - \varphi_{\ell(k)}(x^{k+1})) = -m(f(x^k) - f(x^{k+1})) - m\xi_{k+1}, \tag{41}$$

where $\xi_{k+1} = f(x^{k+1}) - \varphi_{\ell(k)}(x^{k+1}) \ge 0$. This inequality is equivalent to

$$\xi_{k+1} \le \left(\frac{1}{m} - 1\right)\left(f(x^k) - f(x^{k+1})\right). \tag{42}$$

Summing (42) over $k = 0, 1, \ldots, l$ one has $\sum_{k=0}^{l} \xi_{k+1} \le (\frac{1}{m} - 1)(f(x^0) - f(x^{l+1})) \le (\frac{1}{m} - 1)(f(x^0) - f^*)$. Letting $l \to +\infty$, we obtain that $\sum_{k=0}^{\infty} \xi_{k+1} < \infty$.

(ii) From Theorem 4.2-(i), it follows that

$$\tilde{g}^k = -\gamma_k \mathbb{A}^\top \mathcal{Q}^{-1}_{w(x^k)} \mathbb{A}(x^{k+1} - x^k) \in \partial_{\xi_{k+1}} f(x^{k+1}). \tag{43}$$

This means that the sequence $\{x^k\}$ satisfies the conditions of IPAVM algorithm with $\varepsilon_k = \xi_{k+1}$ and $\zeta^{k+1} = 0$ for all $k \in \mathbb{N}$ (see Section 3). Then, (43) implies that

$$f(x^k) + \xi_{k+1} \ge f(x^{k+1}) + \gamma_k \|x^{k+1} - x^k\|^2_{\mathcal{Q}_{w(x^k)}}.$$

Hence, following the same steps of Proposition 3.2 the two first results in (ii) are obtained. For the other part of (ii), we note (41) implies that $f(x^{k+1}) \le f(x^k)$, that is, $\{f(x^k)\}$ is a sequence decreasing. Since $\{f(x^k)\}$ is bounded below by f_*, it converges to some $l \in \mathbb{R}$. Finally, the part (iii) it follows from Proposition 3.3 via suitable identifications. □

5. Numerical experiences

In order to assess from a practical point of view the algorithm IPBAVM described in the previous sections, we have coded it in MATLAB 7.8, Release 2009b and run it on a set of well-known problems in convex nonsmooth optimization with nonnegative and second-order cone constraints, on an IMAC with an Intel Pentium Core i5 CPU 2.7 GHz processor and 8GB of RAM, running OS X operating system. The starting point for this code is a version of the proximal bundle algorithm provided by Professor Claudia Sagastizábal, which solves problems of the type $\min_{x \in \mathbb{R}^n} h \circ c$ with $h : \mathbb{R}^m \to \mathbb{R}$ a positively homogeneous (degree 1) convex function and $c : \mathbb{R}^n \to \mathbb{R}^m$ a smooth mapping (see [45] for more details). Note that Sagastizábal's code does not solve directly the test problems considered in this article.

Additionally, we use CVX (available in http://cvxr.com/cvx) with its default solver (SDPT3) and settings for solving each test problem considered. This will enable us to compute, for instance, the relative error between the computed value at the final iteration by our approach and the obtained by CVX.

Next, we list a set of benchmark examples (see [29,38,46]) with additional constraints which will be used to test our algorithm IPBAVM. This examples have the following form

$$\min_{x \in \mathbb{R}^n} f(x); \quad w(x) \in \mathcal{K},$$

where

- CB2: $f(x) = \max\{x_1^2 + x_2^4, (2 - x_1)^2 + (2 - x_2)^2, 2\exp(x_2 - x_1)\}$
- QL: $f(x) = \max\{x_1^2 + x_2^2, x_1^2 + x_2^2 + 10(-4x_1 - x_2 + 4), x_1^2 + x_2^2 + 10(-x_1 - 2x_2 + 6)\}$
- EVD2: $f(x) = \max\{x_1^2 + x_2^2 + x_3^2 - 1, x_1^2 + x_2^2 + (x_3 - 2)^2, x_1 + x_2 + x_3 - 1, x_1 + x_2 - x_3 + 1, 2x_1^4 + 6x_2^2 + 2(5x_3 - x_1 + 1)^2, x_1^2 - 9x_3\}$
- Mifflin 2: $f(x) = -x_1 + 2(x_1^2 + x_2^2 - 1) + 1.75|x_1^2 + x_2^2 - 1|$
- Rosen-Suzuki (R-S): $f(x) = \max_{1\le j\le 4} f_j(x)$ with

$$\begin{aligned}
f_1(x) &= x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4,\\
f_2(x) &= f_1(x) + 10(x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8),\\
f_3(x) &= f_1(x) + 10(x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10),\\
f_4(x) &= f_1(x) + 10(2x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5)
\end{aligned}$$

- MaxQuad (MQ$_L$): $f(x) = \max_{1\le j\le L}\{\langle A^j x, x\rangle - \langle b^j, x\rangle + c^j\}$, where A^j is a 10×10 symmetric matrix defined by

$$A^j_{ik} = \exp\left(\frac{i}{k}\right)\cos(ik)\sin(ij), \quad i < k \text{ and } A^j_{ii} = \frac{i}{n}|\sin(j)| + \sum_{l\neq i}|A^j_{il}|,$$

  b^j is a vector in $\mathbb{R}^{10}$ whose components are $b^j_i = \exp(i/k)\sin(ij)$, and c^j is a scalar defined by $c^j = 1$, for $j = 1, \dots, L$.

On the other hand, we recall the sufficient condition on the regularization parameter to have interior iterates: $\gamma_k > \|\mathbb{A}^{-1}\|\boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})^{1/2}[\|g^k\| + \delta_k]$ (cf. Proposition 8). In practical computations, we have noticed that this condition can be weakened to speed up convergence. In fact, we implemented the following relaxed version for the parameter:

$$\gamma_k(\theta) = \frac{1}{2^\theta}\left(\|\mathbb{A}^{-1}\|\boldsymbol{\lambda}_{\max}(\mathcal{Q}_{w(x^k)})^{1/2}[\|g^k\| + \delta_k]\right), \quad 0 \le \theta \le \theta_{\max}, \tag{44}$$

and we denote by $y(\theta) = x^k - \gamma_k(\theta)^{-1}(\mathbb{A}^\top \mathcal{Q}^{-1}_{w(x^k)}\mathbb{A})^{-1}\tilde{g}^\ell$. Then we set $y^{\ell+1} = y(\theta^*_{\ell k})$, where $\theta^*_{\ell k} = \max\{0, \dots, \theta_{\max} : y(\theta) \in \mathcal{F}^0\}$.

5.1. Test problems on the nonnegative orthant

In this case, $\mathcal{K} = \mathbb{R}^m_+$. Next, we indicate the dimensions n and m, the feasible starting point x^0 and describe the constraints of each test problem.

- *Test problem 1* (CB2): $n = 2$, $m = 2$, $x^0 = (0.5, 1)^\top$, $w_1(x) = 2x_1 + x_2 - 1$, $w_2(x) = -3x_1 + 4x_2 + 6$.
- *Test problem 2* (QL): $n = 2, m = 3, x^0 = (0.5, 2)^\top, w_1(x) = 2x_1+x_2-1, w_2(x) = -3x_1+4x_2+6$, $w_3(x) = -x_1 - 2x_2 + 14$.
- *Test problem 3* (EVD2): $n = 3$, $m = 3$, $x^0 = (1, 1, 0)^\top$, $w_1(x) = 2x_1 + 3x_2 + x_3 - 4$, $w_2(x) = -4x_1 + 6x_2 + 2x_3 + 8$, $w_3(x) = 5x_1 - 4x_2 - 3x_3 + 10$.
- *Test problem 4* (Mifflin 2): $m = 2$, $m = 2$, $x^0 = (0, 1)^\top$, $w_1(x) = x_1 + x_2 - 0.5$, $w_2(x) = -3x_1 - x_2 + 2.5$.
- *Test problem 5* (R-S): $n = 4$, $m = 4$, $x^0 = (0, 1, 1, 2)^\top$, $w_1(x) = 3x_1 + 2x_2 + 4x_4 - 9$, $w_2(x) = -2x_1+5x_3+6x_4+6$, $w_3(x) = 4x_1-3x_2-4x_3+x_4+10$, $w_4(x) = x_1-x_2+4x_3-2x_4+5$.
- *Test problem 6* (MQ$_L$): $n = 10$, $m = 11$, $x^0 = 0.004$, $w_i(x) = -|x_i| + 0.05$, for $i = 1, \dots, 10$, and $w_{11}(x) = -\sum_{i=1}^n x_i + 0.05$.

Table 1. Average CPU time in seconds used by SDPT3 under CVX.

Problem	CB2	QL	EVD2	Mifflin 2	R-S	MQ_{10}	MQ_{25}	MQ_{50}
$\mathbf{CPU}_{cvx}$	0″.55	0″.68	0″.91	0″.55	1″.45	0″.95	1″.45	2″.24
f_{cvx}	1.9522	7.2	4.92955	−0.943649	−32.2867	0.940517	1	1

Table 2. Numerical results for bundle algorithm on the nonnegative orthant.

	$Tol = 10^{-2}$			
Problem	**NIG**	**NS**	**CPU**	E_r
CB2	12	3	0″.62	5.1854e-5
QL	11	3	0″.56	0.001270
EVD2	17	5	0″.67	0.004821
Mifflin 2	7	4	0″.61	0.011279
R-S	22	4	0″.81	0.011135
MQ_{10}	10	6	0″.63	0.054649
MQ_{25}	15	6	1″.17	0.034605
MQ_{50}	17	7	1″.39	0.032555

Table 3. Numerical results for bundle algorithm on the nonnegative orthant.

	$Tol = 10^{-3}$				$Tol = 10^{-4}$			
Problem	**NIG**	**NS**	**CPU**	E_r	**NIG**	**NS**	**CPU**	E_r
CB2	12	3	0″.62	5.1854e-5	17	4	0″.66	8.0319e-6
QL	13	4	0″.60	3.3796e-5	17	5	0″.64	1.3198e-6
EVD2	25	12	0″.80	0.001126	147	128	1″.35	1.1935e-4
Mifflin 2	11	6	0″.73	9.5005e-4	13	8	0″.76	2.1037e-4
R-S	24	5	0″.85	5.2895e-4	40	19	0″.91	1.0100e-4
MQ_{10}	15	10	0″.64	0.034122	67	48	1″.43	0.005235
MQ_{25}	60	23	1″.68	5.0076e-6	60	23	1″.68	5.0076e-6
MQ_{50}	51	22	1″.98	2.8918e-5	54	23	2″.06	2.0353e-6

Table 1 shows the average CPU time (run five times) in seconds required by SDPT3 under CVX to solve approximately each test problem, and the respective optimal function value (denoted by f_{cvx}).

Note that when increasing the value of L in the test problems MQ_L, CVX's performance deteriorates as the CPU time increases.

Stopping rule and numerical results

The iterates stop when $|w(x^k)^\top s^k| \leq Tol$, where s^k denotes the dual solution at iteration k, and Tol a given tolerance. The obtained numerical results are listed in Tables 2 and 3 for $Tol = 10^{-2}$ and $Tol = 10^{-3}, 10^{-4}$, respectively.

Here, **NIG** represents the number of general iterations, **NS** the number of serious steps, **CPU** represents the average CPU time in seconds (run five times) for solving each problem, and

$$E_r = \frac{|f_{ipbavm} - f_{cvx}|}{|f_{cvx}|}$$

is the relative error at the final iteration, where f_{ipbavm} is the computed objective value at the final iteration of our algorithm IPBAVM.

From Tables 2 and 3, we can observe that output solutions obtained by our algorithm are optimal when compared with the benchmark given by CVX. For $Tol = 10^{-2}$, the relative error in the first example is close to 0.005%, in the next two examples the relative error is close to 0.1% and 0.5%,

Table 4. Average CPU time in seconds used by SDPT3 under CVX.

Problem	CB2	QL	EVD2	Mifflin 2	R-S	MQ_{10}	MQ_{25}	MQ_{50}
$\mathbf{CPU}_{cvx}$	0″.65	0″.71	0″.98	0″.64	1″.63	0″.83	1″.38	2″.22
f_{cvx}	1.95222	7.57813	3.59408	19.5995	−23.2431	0.960938	1	1

respectively, while in the next two examples is close to 1%. In the sixth example, the relative error is close to 5.5%. In the last two examples the relative error is closed to 0.03%.

For $Tol = 10^{-3}$, the relative error in the two first examples are close to 0.005%, in the third example the relative error is close to 0.1%, while in the next two examples are close to 0.09% and 0.05%. In the sixth example the relative error is close to 3.4%. In the last two examples the relative error is close to 0.001% and 0.003%, respectively.

For $Tol = 10^{-4}$, the relative error in the two first examples are close to 0.001%, while in the next three examples are close to 0.01%. In MQ_{10} the relative error is close to 0.5%, while in the last two examples the relative error is between 0.0002% and 0.0005%. Moreover, in all the test problems, we need few iterations for obtaining convergence.

Note that our algorithm is faster than CVX for QL, R-S and MQ_{50} problems.

5.2. Test problems on second-order cones

Here $\mathcal{K} = \mathcal{L}_+^{m_1} \times \cdots \times \mathcal{L}_+^{m_J}$. Next, we indicate the dimension n, number of cones J, dimension of cones m_j and the starting point x^0 of each test problem.

- *Test problem 1* (CB2): $n = 2, J = 1, m = 2, x^0 = (2, 1)^\top$, and $w(x) = x$.
- *Test problem 2* (QL): $n = 2, J = 2, m_1 = m_2 = 2, x^0 = (2, 1)^\top$, $w_1(x) = (x_1 + x_2, -x_1 + x_2)^\top$, and $w_2(x) = (2x_1 + x_2 - 1, -x_1 + 3x_2)^\top$.
- *Test problem 3* (EVD2): $n = 3, J = 2, m_1 = 2, m_2 = 3, x^0 = (1.8860, -0.1890, -0.4081)^\top$, $w_1(x) = (4x_1 + 6x_2 + 3x_3 - 1, -x_1 + 7x_2 - 5x_3 + 2)^\top$, and $w_2(x) = x$.
- *Test problem 4* (Mifflin 2): $n = 2, J = 2, m_1 = m_2 = 2, x^0 = (4, 4)^\top$, $w_1(x) = (5x_1 - 2.5, -3x_1 + 4x_2 + 1.5)^\top$, and $w_2(x) = (11x_1 - 22, -13x_1 + 4x_2 + 42)^\top$.
- *Test problem 5* (R-S): $n = 4, J = 2, m_1 = 3, m_2 = 4, x^0 = (3, 1, 0, 0)^\top$, $w_1(x) = (2x_1 + 3x_2 - 2x_4 - 1, x_1 + 4x_2 - 6x_3 + 5x_4, -x_1 + 8x_3 + 7x_4 + 2)^\top$, and $w_2(x) = x$.
- *Test problem 6* (MQ_L): $n = 10, J = 1, m = 10, x^0 = (2, \omega/\|\omega\|)^\top$ with $\omega \in \mathbb{R}^9$ generated randomly by Matlab's **randn.m** and $w(x) = x$.

In Table 4, we show the average CPU time (run five times) in seconds used by applying SDPT3 under CVX to these test problems.

From this Table, we can note that when increasing the value of L in the test problem MQ_L, the CPU time also increases.

Stopping rule and numerical results

As stopping rule we take

$$\max\left\{\max_{j=1,\dots,J} |\min\{\lambda_1(s_k^j), 0\}|, \max_{j=1,\dots,J}\left\{|w^j(x^k)^\top s_k^j|\right\}\right\} \le Tol,$$

where s^j denotes the dual solution. The last one is due to Theorem 4.4. The numerical results are listed in Table 5, where **CS** is the value of $\max_{j=1,\dots,J}\{|w^j(x^k)^\top s_k^j|\}$ at the final iteration.

From Table 5, we can observe that output solutions obtained by our algorithm IPBAVM are optimal when compared with the benchmark given by CVX. Moreover, for five test problems, we need few iterations for obtaining convergence. The relative error in all the test problems is between

Table 5. Numerical results for bundle algorithm on the second-order cone.

	$Tol = 10^{-2}$				
Problem	NIG	NS	CS	CPU	E_r
CB2	15	3	3.4490e-4	0″.28	7.8388e-7
QL	65	53	9.9698e-5	0″.68	1.3153e-4
EVD2	243	235	9.9773e-5	1″.41	2.6819e-4
Mifflin 2	256	256	1.0631e-5	0″.85	1.0387e-5
R-S	1976	1974	9.9962e-4	5″.82	8.5994e-5
MQ_{10}	47	12	9.3717e-5	0″.51	9.5383e-5
MQ_{25}	27	10	9.3124e-5	0″.98	1.2021e-4
MQ_{50}	27	10	9.3124e-5	1″.22	1.2020e-4

0.01% and 0.008%. For the CB2, QL, R-S and MQ_L (for $L = 10, 25, 50$) problems, IPBAVM is faster than CVX. On the other, these results do not vary when we consider as tolerance $Tol = 10^{-3}, 10^{-4}$.

6. Concluding remarks

In this work, we have proposed an interior proximal-type algorithm with variable metric for solving convex SCP problems. The variable metric is induced by a class of positive definite operators $\mathcal{Q}$. defined on the Euclidean Jordan algebra $\mathbb{V}$. With the introduction of this metric, we showed that the proposed algorithm is well defined, the iterates belong to the interior of the feasible set and we establish some properties of convergence. Moreover, we have adapted and implemented the bundle algorithm for solving nonsmooth convex SCP problems defined on two instances of symmetric cones: the nonnegative orthant and the second-order cone. Then, we have applying these implementations to some well-known test problems, obtaining good results. From the numerical results, we can see that the output solutions obtained by our algorithm are optimal when compared with the benchmark given by CVX. Moreover, we observe that our algorithm, in some test problems, is faster in terms of CPU time than CVX.

Acknowledgements

The authors wish to thank Professor Claudia Sagastizábal who provided us the codes of the bundle algorithm described in [45].

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by FONDECYT under [grant number 1160894], [grant number 1130176]; Millennium Scientific Institute on Complex Engineering Systems [grant number ICM: P-05-004-F], (CONICYT: FBO16). Finally, we thank the anonymous referees for their careful reading and helpful suggestions that improved the paper greatly.

References

[1] Alvarez F, López J, Ramírez HC. Interior proximal algorithm with variable metric for second-order cone programming: applications to structural optimization and support vector machines. Optim. Methods Softw. 2010;25:859–881.
[2] Alizadeh F, Goldfarb D. Second-order cone programming. Math. Program. 2003;95:3–51.
[3] Lobo MS, Vandenberghe L, Boyd S, et al. Applications of second-order cone programming. Linear Algebra Appl. 1998;284:193–228.

[4] Kojima M, Shida M, Shindoh S. Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. Math. Progr. 1998;80:129–160.
[5] Valkonen T. A method for weighted projections to the positive definite cone. Optimization. 2015;64:2253–2275.
[6] Vandenberghe L, Boyd S. Semidefinite programming. SIAM Rev. 1996;38:49–95.
[7] Faybusovich L. Euclidean Jordan algebras and interior point algorithm. Positivity. 1997;1:331–357.
[8] Faybusovich L, Lu Y. Jordan-algebraic aspects of nonconvex optimization over symmetric cone. Appl. Math. Optim. 2006;53:67–77.
[9] Gu G, Zangiabadi M, Roos C. Full Nesterov–Todd step infeasible interior-point method for symmetric optimization. European J. Oper. Res. 2011;214:473–484.
[10] Lu Y, Yuan YX. An interior-point trust-region algorithm for general symmetric cone programming. SIAM J. Optim. 2007;18:65–86.
[11] Schmieta SH, Alizadeh F. Extension of primal-dual interior point algorithms to symmetric cones. Math. Program. Ser. A. 2003;96:409–438.
[12] Wang GQ, Bai YQ. A new full Nesterov-Todd step primal-dual path- following interior-point algorithm for symmetric optimization. J. Optim. Theory Appl. 2012;154:966–985.
[13] Bai YQ, Zhang LP. A full-Newton step interior-point algorithm for symmetric cone convex quadratic optimization. J. Ind. Manag. Optim. 2011;7:891–906.
[14] Wang GQ, Zhang ZH, Zhu DT. On extending primal-dual interior-point method for linear optimization to convex quadratic symmetric cone optimization. Numer. Funct. Anal. Optim. 2013;34:576–603.
[15] Wang G, Yu C, Teo K. A new full Nesterov–Todd step feasible interior-point method for convex quadratic symmetric cone optimization. Appl. Math. Comp. 2013;221:329–343.
[16] Roos C, Terlaky T, Vial JPh. Theory and algorithms for linear optimization: an interior-point approach. Chichester: Wiley; 1997.
[17] Wang GQ, Kong LC, Tao JY, et al. Improved complexity analysis of full Nesterov–Todd step interior-point methods for semidefinite optimization. J. Optim. Theory Appl. 2015;166:588–604.
[18] Darvay Zs. New interior point algorithms in linear programming. Adv. Model. Optim. 2003;5:51–92.
[19] Martinet B. Perturbation des methodes d'optimisation, application. RAIRO Anal. Numer. 1978;12:153–171.
[20] Moreau JJ. Proximité et dualité dans um espace Hilbertien. Bull. Soc. Math. France. 1965;93:273–299.
[21] Rockafellar RT. Monotone operators and the proximal point algorithm. SIAM J. Control Optim. 1976;14:877–898.
[22] Auslender A, Teboulle M. Interior gradient and proximal methods for convex and conic optimization. SIAM J. Optim. 2006;16:697–725.
[23] Censor Y, Zenios SA. The proximal minimization algorithm with D-function. J. Optim. Theory Appl. 1992;73:451–464.
[24] Chen G, Teboulle M. Convergence analysis of proximal-like minimization algorithm using Bregman functions. SIAM J. Optim. 1993;3:538–543.
[25] Iusem AN, Teboulle M. On the convergence rate of entropic proximal optimization methods. Comp. Appl. Math. 1993;12:153–168.
[26] Shor N. Minimization methods for non-differentiable functions. Berlin: Springer-Verlag; 1985.
[27] Kelley JE. The cutting-plane method for solving convex programs. J. SIAM. 1960;8:703–712.
[28] Sharifi-Mokhtarian F, Goffin JL. A nonlinear analytic center cutting plane method for a class of convex programming problems. SIAM J. Optim. 1998;8:1108–1131.
[29] Bonnans JF, Gilbert JCh, Lemaréchal C, et al. Numerical optimization: theoretical and practical aspects, universitext. Berlin: Springer-Verlag; 2003.
[30] Correa R, Lemarèchal C. Convergence of some algorithms for convex minimization. Math. Program. 1993;62:261–275.
[31] Lemarèchal C, Sagastizábal C. Variable metric bundle methods: from conceptual to implementable forms. Math. Program. 1997;76:393–410.
[32] Frangioni A. Generalized bundle methods. SIAM J. Optim. 2002;13:117–156.
[33] Kiwiel KC. A constraint linearization method for nondifferentiable convex minimization. Numer. Math. 1987;51:395–414.
[34] Kiwiel KC. An exact penalty function algorithm for nonsmooth convex constrained minimization problems. IMA J. Numer. Anal. 1985;5:111–119.
[35] Kiwiel KC. Exact penalty functions in proximal bundle methods for constrained convex nondifferentiable minimization. Math. Program. 1991;52:285–302.
[36] Auslender A, Teboulle M. Interior gradient and epsilon-subgradient descent methods for constrained convex minimization. Math. Oper. Res. 2004;29:1–26.
[37] Kiwiel KC. A bundle Bregman proximal method for nondifferentiable convex minimization. Math. Program. 1999;85:241–258.
[38] Sagastizábal C, Solodov M. An infeasible bundle method for nonsmooth convex constrained optimization without a penalty function or a filter. SIAM J. Optim. 2005;16:146–169.

[39] Chua CB, Lin H. An inexact proximal bundle method with applications to convex conic programming. Technical report. Singapore and PR China: Nanyang Technological University and Fujian Normal University; 2013.
[40] Rockafellar RT. Convex analysis. Princeton (NJ): Princeton University Press; 1970.
[41] Faraut J, Korányi A. Analysis on symmetric cones. Oxford: Clarendon Press; 1994.
[42] Vinberg E. The theory of homogeneous cones. Trans. Moscow Math. 1965;12:340–403.
[43] Gowda MS, Sznajder R, Tao J. Some P-properties for linear transformations on Euclidean Jordan algebras. Linear Algebra Appl. 2004;393:203–232.
[44] Iusem AN, Svaiter BF, Teboulle M. Multiplicative interior gradient methods for minimization over the nonnegative orthant. SIAM J. Control Optim. 1996;34:389–406.
[45] Sagastizábal C. Composite proximal bundle method. Math. Program. 2013;140:189–233.
[46] Lukšan L. 2000. Test problems for nonsmooth unconstrained and linearly constrained optimization. Technical Report 798, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague.