J Optim Theory Appl (2008) 139: 227–242
DOI 10.1007/s10957-008-9422-2

Central Paths in Semidefinite Programming, Generalized Proximal-Point Method and Cauchy Trajectories in Riemannian Manifolds

J.X. da Cruz Neto · O.P. Ferreira · P.R. Oliveira · R.C.M. Silva

Published online: 25 April 2008

Abstract The relationships among the central path in the context of semidefinite programming, generalized proximal-point method and Cauchy trajectory in a Riemannian manifolds is studied in this paper. First, it is proved that the central path associated to a general function is well defined. The convergence and characterization of its limit point is established for functions satisfying a certain continuity property. Also, the generalized proximal-point method is considered and it is proved that the correspondingly generated sequence is contained in the central path. As a consequence, both converge to the same point. Finally, it is proved that the central path coincides with the Cauchy trajectory in a Riemannian manifold.

Keywords Central path · Generalized proximal-point methods · Cauchy trajectory · Semidefinite programming · Riemannian manifolds

---

Communicated by T. Rapcsák.

This work was supported in part by CNPq Grant 302618/2005-8, by PRONEX(CNPq), CAPES-PICDT and FUNAPE/UFG.

J.X. da Cruz Neto
DM, Universidade Federal do Piauí, Teresina, PI 64049-500, Brazil
e-mail: jxavier@ufpi.br

O.P. Ferreira
IME, Universidade Federal de Goiás, Goiânia, GO 74001-970, Brazil
e-mail: orizon@mat.ufg.br

P.R. Oliveira
COPPE-Sistemas, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ 21945-970, Brazil
e-mail: poliveir@cos.ufrj.br

R.C.M. Silva (✉)
DM, ICE, Universidade Federal de Amazonas, Manaus, AM 69077-000, Brazil
e-mail: rmesquita@ufam.edu.br

1 Introduction

The extension of concepts and techniques from linear programming to semidefinite programming became attractive after the seminal works due to Alizadeh [1] and Nesterov and Nemirovski [2]. It is well known that the concept of central path, with respect to log barrier function, is very important in several subjects including linear programming and semidefinite programming, see for example Todd [3]. The central path for semidefinite programming problems converges, see Halická et al. [4]. More generally, Graña Drummond and Peterzil [5] established its convergence for analytic convex nonlinear semidefinite programming problems. However, the central path does not converges to the analytic center of the solution set, see Halická et al. [4]. Partial characterizations of the limit point have been given by Halická et al. [6].

Several generalizations of the classical proximal-point method studied in Rockafellar [7] have been considered, due to the important role they play in the development of Augmented Lagrangian algorithms. Early works include Chen and Teboulle [8], Eckstein [9], Iusem [10] and Doljansky and Teboulle [11] who introduced a generalized proximal method for unconstrained convex semidefinite programming problems.

Extensions of concepts and techniques from Euclidean space to Riemannian manifold are natural. It has been done frequently in the last few years, with theoretical objectives and also in order to obtain effective algorithms of optimization on Riemannian manifold setting. Several works dealing with this issue include Rapcsák and Thang [12], Rapcsák [13], Ferreira and Oliveira [14] and Nesterov and Todd [15]. A couple of paper have dealt with the behavior of the Cauchy trajectories in Riemannian manifolds, see for example Alvarez et al. [16] and its references.

The central path with respect to general barrier function, for monotone variational inequality problem, has been considered by Iusem et al. [17] and its well definiteness and convergence properties were obtained. Characterizations of the limit point for some specific problems including linear programming were given, i.e., it was proved that the central path converges to the analytic center of the solution set. Also, Iusem et al. [17] provided a connection among central path, generalized proximal-point sequence and Cauchy trajectory (or gradient trajectories) in Riemannian manifold. It was showed that in some cases, including linear programming, these three concepts are in a certain way equivalent. In particular, this relationship allowed to show that the generalized proximal-point sequence converges to the analytic center of the solution set.

In this paper, we will prove the equivalence among three concepts, namely, central path, generalized proximal-point sequence and Cauchy trajectory in Riemannian manifold, in the context of semidefinite programming. The results obtained are natural extensions of the results of Iusem et al. [17]. We begin by studying the central path for semidefinite programming problems associated to the general function. By assuming that this function satisfies some specific properties, we prove that the central path is well defined. The convergence and characterization is established for functions satisfying a certain continuity property, i.e, we prove that the central path converges to the analytic center of the solution set. After the study of the central path we obtain its equivalence with the generalized proximal-point sequence and Cauchy trajectory in Riemannian manifold.

The organization of our paper is as follows. In Sect. 1.1, we list some basic notation and terminology used in our presentation. In Sect. 2, we describe the semidefinite programming problem and the basic assumptions that will be used throughout the paper. In Sect. 3, we introduce some assumptions in order to guarantee well definiteness of the central path and establish some results about it. In Sect. 4, we describe the proximal-point method and establish its connection with the central path. In Sect. 5, we present the relationship among Cauchy trajectory in Riemannian manifold and central path.

1.1 Notation and Terminology

The following notations and results of matrix analysis are used throughout our presentation, they can be found in Horn and Johnson [18]. $\mathbb{R}^n$ denotes the n-dimensional Euclidean space.

$$\mathbb{R}^n_+ = \{(x_1, \dots, x_n) \in \mathbb{R}^n;\ x_i \geq 0,\ \forall i = 1, \dots, n\}$$

and

$$\mathbb{R}^n_{++} = \{(x_1, \dots, x_n) \in \mathbb{R}^n;\ x_i > 0,\ \forall i = 1, \dots, n\}$$

denote the nonnegative and positive orthants, respectively. The set of all matrices $n \times m$ is denoted by $\mathbb{R}^{n\times m}$. The (i, j)th entry of a matrix $X \in \mathbb{R}^{n\times m}$ is denoted by X_{ij} and the jth column is denoted by X_j. The transpose of $X \in \mathbb{R}^{n\times m}$ is denoted by X^T. The set of all symmetric $n \times n$ matrices is denoted by S^n. The cone of positive semidefinite (resp., definite) $n \times n$ symmetric matrices is denoted by S^n_+ (resp., S^n_{++}) and ∂S^n_+ denotes the boundary of S^n_+. The trace of a matrix $X \in \mathbb{R}^{n\times n}$ is denoted by

$$\operatorname{tr} X \equiv \sum_{i=1}^{p} X_{ii}.$$

Given X and Y in $\mathbb{R}^{n\times m}$, the inner product between them is defined as

$$\langle X, Y\rangle \equiv \operatorname{tr} X^T Y = \sum_{i=1, j=1}^{n,m} X_{ij} Y_{ij}.$$

The Frobenius norm of the matrix X is defined as

$$\|X\| \equiv (\langle X, X\rangle)^{1/2}.$$

The image (or range) space and the null space of a linear operator $\mathcal{A}$ will be denoted by $\mathrm{Im}(\mathcal{A})$ and $\mathrm{Null}(\mathcal{A})$, respectively; the dimension of the subspace $\mathrm{Im}(\mathcal{A})$, referred to as the rank of $\mathcal{A}$; will be denoted by $\mathrm{rank}(\mathcal{A})$. Given a linear operator $\mathcal{A} : E \to F$ between two finite-dimensional inner product spaces $(E, \langle\cdot,\cdot\rangle_E)$ and $(F, \langle\cdot,\cdot\rangle_F)$, its *adjoint* is the unique operator $\mathcal{A}^* : F \to E$ satisfying

$$\langle \mathcal{A}(u), v\rangle_F = \langle u, \mathcal{A}^*(v)\rangle_E, \quad \text{for all } u \in E \text{ and } v \in F.$$

Given $A_1, \dots, A_m \in S^n$, define the linear application $\mathcal{A} : S^n \to \mathbb{R}^m$ by

$$\mathcal{A}X = (\langle A_1, X\rangle, \dots, \langle A_m, X\rangle)^T. \tag{1}$$

Note that the adjoint application $\mathcal{A}^* : \mathbb{R}^m \to S^n$ of $\mathcal{A}$ is given by

$$\mathcal{A}^* v = \sum_{i=1}^{m} v_i A_i.$$

Let $\lambda(X) = (\lambda_1(X), \dots, \lambda_n(X))^T$ denote the vector of eigenvalues of an $n \times n$ matrix X. We assume that the eigenvalues are ordered, e.g., $\lambda_1(X) \geq \lambda_2(X) \geq \cdots \geq \lambda_n(X)$.

Lemma 1.1 *For any* $X, Y \in S^n$, $\mathrm{tr}(XY) \leq \lambda(X)^T \lambda(Y)$.

2 Preliminaries

In this section, we describe our problem and assumptions on it that will be used throughout the paper. Consider the semidefinite programming problem (SDP)

$$\text{(P)} \quad \min\{\langle C, X\rangle : \mathcal{A}X = b, X \succeq 0\},$$

and its associated dual

$$\text{(D)} \quad \max\{b^T y : \mathcal{A}^* y + S = C, S \succeq 0\},$$

where the data consist of $C \in S^n$, $b \in \mathbb{R}^m$ and a linear operator $\mathcal{A} : S^n \to \mathbb{R}^m$, the primal variable is $X \in S^n$, and the dual variable consists of $(S, y) \in S^n \times \mathbb{R}^m$. We write $\mathcal{F}$(P) and $\mathcal{F}$(D) for the sets of feasible solutions to (P) and (D) respectively, and by $\mathcal{F}^0$(P) and $\mathcal{F}^0$(D) strictly feasible solution. We also write $\mathcal{F}^*$(P) and $\mathcal{F}^*$(D) for the sets of optimal solutions of (P) and (D) respectively.

Throughout this paper, we assume that the following two conditions hold without explicitly mentioning them in the statements of our results:

(A1) $\mathcal{A} : S^n \to \mathbb{R}^m$ is a surjective linear operator;
(A2) $\mathcal{F}^0(\text{P}) \neq \emptyset$ and $\mathcal{F}^0(\text{D}) \neq \emptyset$.

A1 is not really crucial for our analysis but it is convenient to ensure that the variables S and y are in one-to-one correspondence. A2 ensures that both (P) and (D) have optimal solutions, the optimal values of (P) and (D) are equal and the solutions sets $\mathcal{F}^*$(P) and sets $\mathcal{F}^*$(D) are bounded (see for example Todd [3]). It is also important to ensure the existence of the central path.

3 Central Paths in Semidefinite Programming

In this section, we describe the central path associated to $\varphi : S^n_{++} \to \mathbb{R}$. Assuming that φ satisfies some assumptions, we prove that the central path is well defined, bounded and converges. Moreover, if φ can be continuously extended to S^n_+ we prove also that the central path converges to the analytic center of the solution set of (P).

Let $\varphi : S^n_{++} \to \mathbb{R}$ be a strictly convex function and C^2. The *central path* to the Problem (P) with respect to φ is the set of points $\{X(\mu) : \mu > 0\}$ defined by

$$X(\mu) = \operatorname*{argmin}_{X \succ 0}\{\langle C, X\rangle + \mu\varphi(X) : \mathcal{A}X = b\}, \quad \mu \in \mathbb{R}_{++}. \tag{2}$$

Some of our results require one of the following assumptions on φ.

(A3) (i) The function φ can be continuously extended to S^n_+ and, for all $\alpha \in \mathbb{R}$, the sublevel set $L_\alpha = \{X \in S^n_+ : \varphi(X) \le \alpha\}$ is bounded.
(ii) For all sequence $\{X_k\} \subset S^n_{++}$ such that $\lim_{k\to\infty} X_k = X \in \partial S^n_+$, there holds $\lim_{k\to\infty}\langle \nabla\varphi(X_k), \tilde{X} - X_k\rangle = -\infty$, for all $\tilde{X} \in S^n_{++}$.

(A4) (i) The function φ goes to $+\infty$ as X goes to the boundary ∂S^n_+ of S^n_+, i.e., $\lim_{X\to\partial S^n_+} \varphi(X) = +\infty$.
(ii) For each $V \in S^n_{++}$ and $\mu > 0$, the function $\phi_{V,\mu} : S^n_{++} \to \mathbb{R}$ defined by $\phi_{V,\mu}(X) = \langle V, X\rangle + \mu\varphi(X)$ satisfies $\lim_{\|X\| \to +\infty} \phi_{V,\mu}(X) = +\infty$.

The above assumptions will be applied to $\varphi : S^n_{++} \to \mathbb{R}$ in the following examples.

Example 3.1 Let $\varphi : S^n_{++} \to \mathbb{R}$ be given by $\varphi(X) = \operatorname{tr}(X\log(X))$. Clearly φ extends continuously to S^n_+ with the convention that $0\log 0 = 0$. The gradient of φ is given by $\nabla\varphi(X) = \log(X) + I$. It is easy to see that the function φ is strictly convex and has a unique minimizer $X^* = e^{-I}$. Therefore, $L_\alpha = \{X \in S^n_+ : \varphi(X) \le \alpha\}$ is bounded. Let us consider $\{X_k\} \subset S^n_{++}$ such that $\lim_{k\to\infty} X_k = X \in \partial S^n_+$ and $\tilde{X} \in S^n_{++}$. So,

$$\begin{aligned}\langle \nabla\varphi(X_k), \tilde{X} - X_k\rangle &= \langle \log(X_k) + I, \tilde{X} - X_k\rangle \\ &= \langle \log(X_k), \tilde{X}\rangle - \langle \log(X_k), X_k\rangle + \langle I, \tilde{X}\rangle - \langle I, X_k\rangle \\ &\le \sum_{i=1}^n \lambda_i(\tilde{X})\lambda_i(\log(X_k)) - \varphi(X_k) + \sum_{i=1}^n \lambda_i(\tilde{X}) - \sum_{i=1}^n \lambda_i(X_k).\end{aligned}$$

Since $\tilde{X} \in S^n_{++}$ and $\lim_{k\to\infty} X_k = X \in \partial S^n_+$, the first term of the right-hand side of the last inequality goes to $-\infty$ as k goes to ∞, and due to the fact that the other ones have a finite limit, we obtain that $\lim_{k\to\infty}\langle \nabla\varphi(X_k), \tilde{X} - X_k\rangle = -\infty$. Hence, φ satisfies A3.

For details on the properties of $\log(X)$, see Horn and Johnson [18].

Example 3.2 Let $\varphi : S^n_{++} \to \mathbb{R}$ be given by $\varphi(X) = -\log\det(X)$. The gradient of φ is given by $\nabla\varphi(X) = -X^{-1}$. It easy to see that φ is strictly convex. So, for each $V \in S^n_{++}$ and $\mu > 0$ the function $\phi_{V,\mu}(X) = \langle V, X\rangle - \mu\log\det(X)$ is also strictly convex. Since $\phi_{V,\mu}$ is convex, we obtain

$$\phi_{V,\mu}(X) \ge \phi_{V,\mu}(2\mu V^{-1}) + \langle \nabla\phi_{V,\mu}(2\mu V^{-1}), X - 2\mu V^{-1}\rangle,$$

or equivalently,

$$\phi_{V,\mu}(X) \geq \langle V, X\rangle/2 + \mu n - \mu \log\det(2\mu V^{-1}).$$

As $V \in S^n_{++}$ we have $\lim_{\|X\|\to+\infty}\langle V, X\rangle = +\infty$ thus, the latter inequality implies $\lim_{\|X\|\to+\infty}\phi_{V,\mu}(X) = +\infty$. Now, $\lim_{X\to\partial S^n_{++}} -\log\det(X) = +\infty$. Therefore, φ satisfies A4.

Example 3.3 Let $\varphi : S^n_{++} \to \mathbb{R}$ be given by $\varphi(X) = \det X^{-\alpha}$, where $\alpha > 0$. It is easy to show that the gradient and the Hessian of φ are given, respectively, by

$$\nabla\varphi(X) = -\alpha \det X^{-\alpha} X^{-1},$$
$$\nabla^2\varphi(X)H = \alpha \det X^{-\alpha}(\alpha\langle X^{-1}, H\rangle X^{-1} + X^{-1}HX^{-1}),$$

where $H \in S^n$. Hence, we obtain

$$\langle\nabla^2\varphi(X)H, H\rangle = \alpha\det X^{-\alpha}(\alpha\langle X^{-1}, H\rangle^2 + \|X^{-1/2}HX^{-1/2}\|^2) > 0,$$

for all $H \neq 0$. So, φ is strictly convex; as a consequence,

$$\phi_{V,\mu}(X) = \langle V, X\rangle + \mu\det X^{-\alpha}$$

is also strictly convex, for all $V \in S^n_{++}$ and $\mu > 0$. Now, since $\mu > 0$, $\det X^{-\alpha} > 0$ and $V \in S^n_{++}$ we have $\lim_{\|X\|\to+\infty}\phi_{V,\mu}(X) \geq \lim_{\|X\|\to+\infty}\langle V, X\rangle = +\infty$. Finally, as $\lim_{X\to\partial S^n_{++}} \det X^{-\alpha} = +\infty$, we have that φ satisfies A4.

Example 3.4 Let $\varphi : S^n_{++} \to \mathbb{R}$ be given by $\varphi(X) = \operatorname{tr} X^{-1}$. So, $\nabla\varphi(X) = -X^{-2}$ and $\nabla^2\varphi(X)H = X^{-2}HX^{-1} + X^{-1}HX^{-2}$, where $H \in S^n$. Then, we obtain

$$\langle\nabla^2\varphi(X)H, H\rangle = \|X^{-1}HX^{-1/2}\|^2 + \|X^{-1/2}HX^{-1}\|^2 > 0,$$

for all $H \neq 0$, which implies that φ is strictly convex. Thus, $\phi_{V,\mu}(X) = \langle V, X\rangle + \mu\operatorname{tr} X^{-1}$ is also strictly convex, for all $V \in S^n_{++}$ and $\mu > 0$. As $\operatorname{tr} X^{-1} > 0$, $\mu > 0$ and $V \in S^n_{++}$ we have $\lim_{\|X\|\to+\infty}\phi_{V,\mu}(X) \geq \lim_{\|X\|\to+\infty}\langle V, X\rangle = +\infty$. Now, since $\lim_{X\to\partial S^n_{++}} \operatorname{tr} X^{-1} = +\infty$, we obtain that φ satisfies A4.

A3 and A4 are important to assure the well definiteness of the central path. Now, we are going to show that A4 implies that $\phi_{V,\mu}$ has compact sublevel sets.

Lemma 3.1 *Under* A4, *the sublevel set* $K_{\alpha,\mu}(V) = \{X \in S^n_{++} : \phi_{V,\mu}(X) \leq \alpha\}$ *is compact for each* $\alpha \in \mathbb{R}$. *As a consequence*, $\phi_{V,\mu}$ *has a minimizer in* S^n_{++}.

Proof Let $\alpha \in \mathbb{R}$. We claim that $K_{\alpha,\mu}(V)$ is bounded. Indeed, assume by contradiction that $K_{\alpha,\mu}(V)$ is unbounded. Then, there exists a sequence $\{X_k\} \subset K_{\alpha,\mu}(V)$ such that $\lim_{k\to+\infty}\|X_k\| = +\infty$. But, A4(ii) implies $\lim_{\|X_k\|\to+\infty}\phi_{V,\mu}(X_k) = +\infty$ which is absurd, since $\phi_{V,\mu}(X_k) \leq \alpha$, for all k. So, we have established the claim. Now we are going to show that $K_{\alpha,\mu}(V)$ is closed. Let $\{X_k\} \subset K_{\alpha,\mu}(V)$ be such that $\lim_{k\to+\infty} X_k = \bar{X}$. Since $\{X_k\} \subset S^n_{++}$, we have two possibilities: $\bar{X} \in S^n_{++}$ or $\bar{X} \in \partial S^n_{++}$, where ∂S^n_{++} denotes the boundary of S^n_{++}. Since $\mu > 0$, A4(i) implies that $\bar{X} \notin \partial S^n_{++}$. So, $\bar{X} \in S^n_{++}$. Now, the continuity of $\phi_{V,\mu}$ in S^n_{++} implies that $\alpha \geq \lim_{k\to+\infty} \phi_{V,\mu}(X_k) = \phi_{V,\mu}(\bar{X})$, i.e., $\bar{X} \in K_{\alpha,\mu}$. Thus, $K_{\alpha,\mu}(V)$ is closed. Therefore $K_{\alpha,\mu}(V)$ is compact, and it is easy to conclude that $\phi_{V,\mu}$ has a minimizer in S^n_{++}. □

Theorem 3.1 *If φ satisfies* A3 *or* A4, *then the central path $\{X(\mu) : \mu > 0\}$ is well defined and is in $\mathcal{F}^0$*(P).

Proof Take $X_0 \in \mathcal{F}^0$(P) and $(y_0, S_0) \in \mathcal{F}^0$(D). For $\mu > 0$, define $\phi_{(S_0,\mu)} : S^n_{++} \to \mathbb{R}$ by

$$\phi_{(S_0,\mu)}(X) = \langle S_0, X\rangle + \mu\varphi(X).$$

First, assume that φ satisfies A3. It easy to see that

$$\langle C, X\rangle = \langle S_0, X\rangle + b^T y_0, \quad \text{for all } X \in \mathcal{F}(\text{P});$$

hence, (2) is equivalent to

$$X(\mu) = \operatorname*{argmin}_{X\succ 0}\{\phi_{(S_0,\mu)}(X) : \mathcal{A}X = b,\ \phi_{(S_0,\mu)}(X) \leq \phi_{(S_0,\mu)}(X_0)\}. \tag{3}$$

Let us consider the sublevel set

$$L_\beta = \{X \in S^n_+ : \phi_{(S_0,\mu)}(X) \leq \beta\},$$

where $\beta = \phi_{(S_0,\mu)}(X_0)$. Note that

$$L_\beta \subset L_\alpha = \{X \in S^n_+ : \varphi(X) \leq \alpha\}, \quad \text{where } \alpha = \beta/\mu.$$

From A3(i) we have that L_α is bounded. So, L_β is also bounded. As $\phi_{(S_0,\mu)}$ is continuous in S^n_+ we have that L_β is compact, which implies that $L_\beta \cap \{X \in S^n_+ : \mathcal{A}X = b\}$ is also compact. Since $\phi_{(S_0,\mu)}$ is strictly convex we have that there exists a unique minimizer $X(\mu) \in \mathcal{F}$(P) and therefore (3) is well defined. Thus, the central path $\{X(\mu) : \mu > 0\}$ is also well defined. Now, we are going to show that $X(\mu) \in \mathcal{F}^0$(P). Assume by contradiction that

$$X(\mu) \in \partial\mathcal{F}(\text{P}) = \{X \in \partial S^n_+ : \mathcal{A}X = b\}.$$

Define the sequence

$$Z_k = (1 - \varepsilon_k)X(\mu) + \varepsilon_k X_0,$$

where $\{\varepsilon_k\}$ is a sequence satisfying $\varepsilon_k \in (0, 1)$ and $\lim_{k\to\infty} \varepsilon_k = 0$. Then, as $X_0 \in \mathcal{F}^0$(P), $X(\mu) \in \partial\mathcal{F}$(P), $\varepsilon_k \in (0, 1)$ and $\mathcal{F}^0$(P) is convex, we conclude that $Z_k \in \mathcal{F}^0$(P) for all $\varepsilon_k \in (0, 1)$. Combining the definitions of $X(\mu)$ and sequence $\{Z_k\}$

with the convexity of φ, we obtain

$$\begin{aligned}0 &\le \langle C, Z_k\rangle + \mu\varphi(Z_k) - \langle C, X(\mu)\rangle - \mu\varphi(X(\mu))\\ &\le \langle C, (Z_k - X(\mu))\rangle + \mu\langle\nabla\varphi(Z_k), Z_k - X(\mu)\rangle\\ &= \varepsilon_k\langle C, (X_0 - X(\mu))\rangle + \mu\frac{\varepsilon_k}{1-\varepsilon_k}\langle\nabla\varphi(Z_k), X_0 - Z_k\rangle.\end{aligned}$$

The latter inequality implies that

$$[(1-\varepsilon_k)/\mu]\langle C, (X(\mu) - X_0)\rangle \le \langle\nabla\varphi(Z_k), X_0 - Z_k\rangle.$$

Since $\lim_{k\to\infty} Z_k = X(\mu) \in \partial\mathcal{F}(\mathrm{P})$, A3(ii) implies that the right-hand side of the above inequality goes to $-\infty$, as k goes to ∞, however the left-hand side of this inequality has a finite limit. Therefore, this contradiction implies that $X(\mu) \in \mathcal{F}^0(\mathrm{P})$.

Finally, assume that φ satisfies A4. Let us consider the sublevel set

$$K_{\alpha,\mu}(S_0) = \{X \in S^n_{++} : \phi_{(S_0,\mu)}(X) \le \alpha\},$$

where $\alpha = \phi_{(S_0,\mu)}(X_0)$. From Lemma 3.1 we have that $K_{\alpha,\mu}(S_0) \cap \{X \in S^n_+ : \mathcal{A}X = b\}$ is compact. So, a similar argument used in the first part permit to conclude that $\{X(\mu) : \mu > 0\}$ is well defined and definition of $K_{\alpha,\mu}(S_0)$ implies that it is in $\mathcal{F}^0(\mathrm{P})$. □

For φ satisfying A3 or A4, the latter theorem implies that $\{X(\mu) : \mu > 0\}$, associated to φ, is well defined and is in $\mathcal{F}^0(\mathrm{P})$. So, for all $\mu > 0$, we have from (2) that

$$\mu\nabla\varphi(X(\mu)) = -C + \mathcal{A}^* y(\mu), \tag{4}$$

for some $y(\mu) \in \mathbb{R}^m$.

Proposition 3.1 *Suppose that φ satisfies* A3 *or* A4. *Then, we have*:

(i) *the function $0 < \mu \mapsto \varphi(X(\mu))$ is nonincreasing*;
(ii) *the set $\{X(\mu) :\ 0 < \mu < \bar{\mu}\}$ is bounded, for each $\bar{\mu} > 0$*;
(iii) *all cluster points of $\{X(\mu) : \mu > 0\}$ are solutions of Problem* (P).

Proof The proof of items (i), (ii), (iii) are similar to the proof of the Proposition 3(i), Proposition 4 and Proposition 5 of Iusem et al. [17], respectively. □

Theorem 3.2 *Assume that φ satisfies* A3. *Let $X^* \in S^n_+$ be the analytic center of $\mathcal{F}^*(\mathrm{P})$ with respect to φ, i.e., the unique solution of the problem*

$$(\mathrm{S}) \quad \min\{\varphi(X) : X \in \mathcal{F}^*(\mathrm{P})\},$$

Then $\lim_{\mu\to+\infty} X(\mu) = X^*$, where $\{X(\mu) : \mu > 0\}$ is the central path with respect to φ.

Proof Take $\bar{X}$ a cluster point of $\{X(\mu) : \mu > 0\}$ and a sequence of positive numbers $\{\mu_k\}$ such that $\lim_{k\to+\infty} \mu_k = 0$ and $\lim_{k\to+\infty} X(\mu_k) = \bar{X}$. Now, from (4) we have $C + \mu_k \nabla\varphi(X(\mu_k)) = \mathcal{A}^* y(\mu_k)$, for some $y(\mu_k) \in \mathbb{R}^m$. So,

$$\langle \mu_k \nabla\varphi(X(\mu_k)), X - X(\mu_k)\rangle = \langle \mathcal{A}^* y(\mu_k) - C, X - X(\mu_k)\rangle,$$

for all $X \in \mathcal{F}^*(\mathrm{P})$. Using the fact that $X - X(\mu_k) \in \mathrm{Null}(\mathcal{A})$, this equality becomes

$$\langle \mu_k \nabla\varphi(X(\mu_k)), X - X(\mu_k)\rangle = -\langle C, X - X(\mu_k)\rangle.$$

Since φ is convex, the above equality implies that

$$\mu_k(\varphi(X(\mu_k)) - \varphi(X)) \leq \langle C, X\rangle - \langle C, X(\mu_k)\rangle.$$

Because $X \in \mathcal{F}^*(\mathrm{P})$ and $\mu_k > 0$, it follows from the latter inequality that $\varphi(X(\mu_k)) \leq \varphi(X)$. Now, as φ is continuous we can take limits, as k goes to $+\infty$, in this inequality to conclude that $\varphi(\bar{X}) \leq \varphi(X)$, for all $X \in \mathcal{F}^*(\mathrm{P})$. Thus, any cluster point $\bar{X}$ of $\{X(\mu) : \mu > 0\}$ is a solution of the problem (S). Therefore, since X^* is the unique solution of (S), the central path converges to it and the theorem is proved. □

Theorem 3.3 *The central path* $\{X(\mu) : \ \mu > 0\}$ *with respect to the function* $\varphi(X) = -\log\det(X)$ *converges*, *as* μ *goes to* 0.

Proof See Halická et al. [4] or Graña Drummond and Peterzil [5]. □

4 Central Paths and Generalized Proximal-Point Methods

In this section, we describe a generalized proximal-point method to solve semidefinite programming problem and present some convergence results for it. It is worthwhile to mention that our goal in this section is to bring to semidefinite programming context the ideas of Iusem et al. [17].

We begin by defining a generalized distance. Let $\varphi : S^n_{++} \to \mathbb{R}$ be a C^2 strictly convex function. Define the generalized distance $D_\varphi : S^n_{++} \times S^n_{++} \to \mathbb{R}$, with respect to φ, as

$$D_\varphi(X, Y) = \varphi(X) - \varphi(Y) - \langle \nabla\varphi(Y), X - Y\rangle.$$

Example 4.1 For the functions of the Examples 3.1–3.4, their correspondent generalized distances are given, respectively, by

(i) $D_\varphi(X, Y) = \mathrm{tr}(X\log X - X\log Y + Y - X)$,
(ii) $D_\varphi(X, Y) = \mathrm{tr}(XY^{-1}) - \log\det(XY^{-1}) - n$,
(iii) $D_\varphi(X, Y) = (\det X)^{-\alpha} + (\det Y)^{-\alpha}(\alpha\,\mathrm{tr}(XY^{-1}) - \alpha n - 1)$, where $\alpha > 0$,
(iv) $D_\varphi(X, Y) = \mathrm{tr}(X^{-1}) - 2\,\mathrm{tr}(Y^{-1}) + \mathrm{tr}(XY^{-2})$.

Remark 4.1 Note that, if φ is C^2, and strictly convex then for each fixed $Y \in S^n_{++}$ the function $D_\varphi(., Y)$ is also C^2 and strictly convex. Moreover, there holds:

(i) if φ satisfies A3, then $D_\varphi(.,Y)$ satisfies A3;
(ii) if φ satisfies A4, then $D_\varphi(.,Y)$ satisfies A4.

Indeed, by assuming that φ satisfies A3, it is immediate to conclude that $D_\varphi(.,Y)$ also can be extended continuously to S^n_+. Now, since $D_\varphi(.,Y)$ has a minimizer and is strictly convex in S^n_{++}, we have that the sublevel set $L_\alpha = \{X \in S^n_+ : D_\varphi(X,Y) \le \alpha\}$ is compact, for all $\alpha \in \mathbb{R}$. Hence $D_\varphi(.,Y)$ satisfies A3(i). Let us consider $\{X_k\} \subset S^n_{++}$ such that $\lim_{k\to\infty} X_k = X \in \partial S^n_+$. So,

$$\langle \nabla D_\varphi(X_k,Y), \tilde{X} - X_k\rangle = \langle \nabla\varphi(X_k), \tilde{X} - X_k\rangle - \langle \nabla\varphi(Y), \tilde{X} - X_k\rangle,$$

for all $\tilde{X} \in S^n_{++}$. By A3(ii) we have that the first term of the right side of last equality goes to $-\infty$ and the second one converges, as k goes to ∞. Therefore, $D_\varphi(.,Y)$ satisfies A3.

Now, assume that φ satisfies A4. Let $V \in S^n_{++}$ and $\mu > 0$. Since $D_\varphi(X,Y) \ge 0$, for $X, Y \in S^n_{++}$, and $V \in S^n_{++}$ it is easy to see that

$$\lim_{\|X\| \to +\infty} (\langle V, X\rangle + \mu D_\varphi(X,Y)) \ge \lim_{\|X\| \to +\infty} \langle V, X\rangle = +\infty.$$

On the other hand, if φ satisfies A4(i), then we have

$$\lim_{X\to\partial S^n_{++}} D_\varphi(X,Y) = \lim_{X\to\partial S^n_{++}} (\varphi(X) - \varphi(Y) - \langle \nabla\varphi(Y), X - Y\rangle) = +\infty.$$

Therefore, $D_\varphi(.,Y)$ satisfies A4, and the statements are proved.

Let $X_0 \in \mathcal{F}^0(\mathrm{P})$. The proximal-point method with generalized distance D_φ, for solving the problem (P), generates a sequence $\{X_k\} \subset S^n_{++}$ with starting point $X_0 \in \mathcal{F}^0(\mathrm{P})$ and

$$X_{k+1} = \operatorname*{argmin}_{X\in S^n_{++}} \{\langle C, X\rangle + \lambda_k D_\varphi(X, X_k) : \mathcal{A}X = b\},$$

where the sequence $\{\lambda_k\} \subset \mathbb{R}_{++}$ satisfies

$$\sum_{k=0}^{\infty} {\lambda_k}^{-1} = +\infty. \tag{5}$$

From now on we refer to the above sequence $\{X_k\}$ as proximal-point sequence with respect to D_φ, associated to $\{\lambda_k\}$ and starting point X_0. If φ satisfies A3 or A4, then Remark 4.1 permits to use a similar argument to prove the well definiteness of the proximal-point sequence $\{X_k\}$, Moreover, $\{X_k\}$ satisfies

$$C + \lambda_k(\nabla\varphi(X_{k+1}) - \nabla\varphi(X_k)) = \mathcal{A}^* z_k, \tag{6}$$

for some sequence $\{z_k\}$ in $\mathbb{R}^m$ and $k = 0, 1, 2, \ldots$.

Let $\{X(\mu) : \mu > 0\}$ be the central path for the Problem (P), with respect to the function $D_\varphi(., X_0)$, i.e.,

$$X(\mu) = \operatorname*{argmin}_{X\in S^n_{++}} \{\langle C, X\rangle + \mu D_\varphi(X, X_0) : \mathcal{A}X = b\},$$

for all $\mu > 0$. If φ satisfies A3 or A4, then from Remark 4.1 and Theorem 3.1, we obtain that the central path $\{X(\mu) : \mu > 0\}$ is well defined and is the unique solution of the system

$$C + \mu(\nabla\varphi(X(\mu)) - \nabla\varphi(X_0)) = \mathcal{A}^* y(\mu), \tag{7}$$

for some path $\{y(\mu) : \mu > 0\}$ in $\mathbb{R}^m$.

Proposition 4.1 *Suppose that φ satisfies* A3 *or* A4. *Then, the central path $\{X(\mu) : \mu > 0\}$, with respect to the function $D_\varphi(., X_0)$, is bounded and all its cluster points are solutions to the Problem* (P). *Moreover, if φ satisfies* A3 *then $X(\mu)$ converges, as μ goes to* 0, *to the analytic center of $\mathcal{F}^*$*(P), *with respect to $D_\varphi(., X_0)$.*

Proof First, note that by Remark 4.1 $D_\varphi(., X_0)$ satisfies the corresponding properties to φ. So, the limitation of $\{X(\mu) : \bar{\mu} > \mu > 0\}$ follows from Proposition 3.1. Moreover, Proposition 3.1 guarantees that all cluster points of $\{X(\mu) : \mu > 0\}$ are solutions to the Problem (P). Finally, if φ satisfies A3, then from Theorem 3.2 we obtain that $\{X(\mu) : \mu > 0\}$ converges to the analytic center of $\mathcal{F}^*$(P) with respect to $D_\varphi(., X_0)$. □

Next we show that the proximal-point sequence $\{X_k\}$ is on the central path. This idea has appeared in Iusem et al. [17], they have proved this connection among central path and generalized proximal-point sequence in some special cases, including linear programming. It is worth pointing out that the next theorem is a natural extension to semidefinite programming of the Theorem 3 of Iusem et al. [17].

Theorem 4.1 *Assume that φ satisfies* A3 *or* A4. *Let $\{X(\mu) : \mu > 0\}$ be the central path with respect to $D_\varphi(., X_0)$ and let $\{X_k\}$ be the proximal-point sequence. If the sequence $\{\mu_k\}$ is defined as*

$$\mu_k = \left(\sum_{j=0}^{k-1} {\lambda_j}^{-1}\right)^{-1}, \quad \textit{for } k = 1, 2, \ldots, \tag{8}$$

then $X_k = X(\mu_k)$, for $k = 1, 2, \ldots$. Moreover, for each positive decreasing sequence $\{\mu_k\}$, there exists a positive sequence $\{\lambda_k\}$ satisfying (5) *such that the proximal sequence $\{X_k\}$ associated to it satisfies $X_k = X(\mu_k)$.*

Proof From (6), we have that X_j satisfies

$$C + \lambda_j(\nabla\varphi(X_{j+1}) - \nabla\varphi(X_j)) = \mathcal{A}^* z_j,$$

or equivalently,

$$C/\lambda_j + \nabla\varphi(X_{j+1}) - \nabla\varphi(X_j) = \mathcal{A}^*(z_j/\lambda_j), \quad \text{for } j = 0, 1, \ldots.$$

Summing this equality from $j = 0$ to $k - 1$, letting μ_k as in (8) and

$$v_k = \mu_k \sum_{j=0}^{k-1} {\lambda_j}^{-1} z_j,$$

we obtain

$$C + \mu_k(\nabla\varphi(X_k) - \nabla\varphi(X_0)) = \mathcal{A}^* v_k, \quad \text{for all } k \geq 1.$$

So, the above equality and (7) implies that $X_k = X(\mu_k)$ and the first part is proved.

For the second part, let $\{X(\mu) : \mu > 0\}$ be the central path and let $\{y(\mu) : \mu > 0\}$ be given by (7). Take a positive decreasing sequence $\{\mu_k\}$ and define the sequences $X_k = X(\mu_k)$ and $y_k = y(\mu_k)$. It follows from (7) that

$$C/\mu_k + \nabla\varphi(X_k) - \nabla\varphi(X_0) = \mathcal{A}^*(y_k/\mu_k).$$

Letting

$$\lambda_k = ((\mu_{k+1})^{-1} - (\mu_k)^{-1})^{-1}$$

and

$$z_k = \lambda_k((\mu_{k+1})^{-1} y_{k+1} - (\mu_k)^{-1} y_k,$$

it is easy to see from last equality that

$$C + \lambda_k(\nabla\varphi(X_{k+1}) - \nabla\varphi(X_k)) = \mathcal{A}^* z_k.$$

Thus, as $\lambda_k = ((\mu_{k+1})^{-1} - (\mu_k)^{-1})^{-1}$ satisfies (5) the result follows from latter equality and (6). □

Theorem 4.2 *Assume that* φ *satisfies* A3 *or* A4. *Let* $\{X_k\}$ *be the proximal-point sequence. Then* $\{X_k\}$ *is bounded and all its cluster points are solutions of the Problem* (P). *Moreover, if* φ *satisfies* A3, *then* $\{X_k\}$ *converges to the analytic center of solution set of the Problem* (P) *with respect to* $D_\varphi(., X_0)$.

Proof If $\{X_k\}$ is the proximal-point sequence, then setting μ_k, as in (8), we obtain from Theorem 4.1 that $X_k = X(\mu_k)$. Since $\lim_{k\to+\infty} \mu_k = 0$, the result follows from Proposition 4.1. □

Lemma 4.1 *Suppose that* φ *satisfies* A3 *or* A4. *Let* $\{X(\mu) : \mu > 0\}$ *be the central path to the Problem* (P) *with respect to* $D_\varphi(., X_0)$. *If* $\nabla\varphi(X_0) \in \operatorname{Im}\mathcal{A}^*$, *then* $\{X(\mu) : \mu > 0\}$ *is also the central path to the Problem* (P) *with respect to* φ.

Proof Take φ satisfying A3 or A4. From Remark 4.1, we have that $D_\varphi(., X_0)$, also satisfies A3 or A4. So, from Theorem 3.1 the central path $\{X(\mu) : \mu > 0\}$ is well defined and satisfies

$$C + \mu(\nabla\varphi(X(\mu)) - \nabla\varphi(X_0)) = \mathcal{A}^*(y(\mu)),$$

for some path $\{y(\mu) : \mu > 0\}$ in $\mathbb{R}^m$. As $\nabla\varphi(X_0) \in \operatorname{Im}\mathcal{A}^*$, there exists $y_0 \in \mathbb{R}^m$ such that $\mathcal{A}^* y_0 = \nabla\varphi(X_0)$; thus, combining the last two equalities, we obtain $C + \mu\nabla\varphi(X(\mu)) = \mathcal{A}^*(y(\mu) + \mu y_0)$. So, $\{X(\mu) : \mu > 0\}$ is a solution to (4), however (4) has unique solution, namely, the central path to the Problem (P) with respect to φ. □

Lemma 4.2 *Let $\{X(\mu) : \mu > 0\}$ be the central path to the Problem* (P) *with respect to $D_\varphi(., X_0)$, where*

$$\varphi(X) = -\log\det(X).$$

If $X_0^{-1} \in \operatorname{Im}\mathcal{A}^*$, then $X(\mu)$ converges as μ goes to 0.

Proof From Example 3.2 we have that $\varphi(X) = -\log\det(X)$ satisfies A4. Since $\nabla\varphi(X_0) = -X_0^{-1} \in \operatorname{Im}\mathcal{A}^*$ we have from Lemma 4.1 that the central paths with respect to $D_\varphi(., X_0)$ and to φ are equal. Therefore, the statement follows from Theorem 3.3. □

Theorem 4.3 *Let $\{X(\mu) : \mu > 0\}$ be the central path with respect to the $D_\varphi(., X_0)$, where $\varphi(X) = -\log\det(X)$. Suppose that $X_0^{-1} \in \operatorname{Im}\mathcal{A}^*$. Then, the proximal-point sequence $\{X_k\}$ with starting point X_0, converges.*

Proof If $\{X_k\}$ is the proximal-point sequence with starting point X_0, then setting μ_k, as in (8), we obtain from Theorem 4.1 that $X_k = X(\mu_k)$. Since $\lim_{k\to+\infty}\mu_k = 0$, the result follows from Lemma 4.2. □

5 Central Paths and Cauchy Trajectories in Riemannian Manifolds

In this section, we are going to prove that the central path, with respect to the function φ for the Problem (P), becomes a Cauchy trajectory on the Riemannian manifold endowed with the metric given by the Hessian of φ. This result extends to semidefinite programming context the corresponding result of linear programming, see Sect. 4 of Iusem et al. [17].

We begin with some basics results of Riemannian geometry. Consider the set of positive definite $n \times n$ symmetric matrices S^n_{++} with its usual differentiable structure and endowed with the Euclidean metric $\langle ., .\rangle$. The tangent space to S^n_{++} at X is given by

$$T_X S^n_{++} = \{Y - X; Y \in S^n\} = S^n.$$

Let $\varphi : S^n_{++} \to \mathbb{R}$ be strictly convex and C^2. Define a new metric in S^n_{++} as

$$\langle U, V\rangle_{\nabla^2\varphi(X)} = \langle \nabla^2\varphi(X)U, V\rangle,$$

so that $M := (S^n_{++}, \nabla^2\varphi)$ is now a Riemannian manifold. The metric of M induces a map $\phi \mapsto \operatorname{grad}\phi$ which associates to each $\phi \in C^1(S^n_{++})$ its *gradient* $\operatorname{grad}\phi \in S^n$ by the rule

$$d\phi_X(V) = \langle \operatorname{grad}\phi(X), V\rangle_{\nabla^2\varphi(X)},$$

where $X \in S^n_{++}$ and $V \in S^n$. It is easy to see that the *gradient vector field* in M is

$$\operatorname{grad}\phi(X) = (\nabla^2\varphi(X))^{-1}\nabla\phi(X),$$

where $\nabla\phi(X)$ is the Euclidean gradient vector field at X, i.e., $\nabla\phi(X)$ is the gradient with respect to the Euclidean metric.

Let $\mathcal{A}: S^n \to \mathbb{R}^m$ be the linear operator as defined in (1). The Assumptions A1 and A2 imply that the set

$$\mathcal{F}^0(\mathrm{P}) = \{X \in S^n : \mathcal{A}X = b,\ X \in S^n_{++}\}$$

is a Riemannian submanifold of M with the induced metric and tangent space at X given by

$$T_X\mathcal{F}^0(\mathrm{P}) = \{V \in S^n; \mathcal{A}V = 0\}.$$

The adjoint operator of $\mathcal{A}$ with respect to the metric of M is $(\nabla^2\varphi(X))^{-1}\mathcal{A}^*$, where $\mathcal{A}^*: \mathbb{R}^m \to S^n$ is the usual adjoint operator of $\mathcal{A}$. In this case, the orthogonal projection $\Pi_X : S^n \to T_X\mathcal{F}^0(\mathrm{P})$ with respect to the metric of M is

$$\Pi_X = I - (\nabla^2\varphi(X))^{-1}\mathcal{A}^*(\mathcal{A}(\nabla^2\varphi(X))^{-1}\mathcal{A}^*)^{-1}\mathcal{A}.$$

The gradient vector field of the function $\phi_{|_{\mathcal{F}^0(\mathrm{P})}} : \mathcal{F}^0(\mathrm{P}) \to \mathbb{R}$, with respect to the metric of M, is given by

$$\operatorname{grad}\phi_{|_{\mathcal{F}^0(\mathrm{P})}} = \Pi \operatorname{grad}\phi,$$

i.e.,

$$\operatorname{grad}\phi_{|_{\mathcal{F}^0(\mathrm{P})}} = (I - (\nabla^2\varphi(X))^{-1}\mathcal{A}^*(\mathcal{A}(\nabla^2\varphi(X))^{-1}\mathcal{A}^*)^{-1}\mathcal{A})(\nabla^2\varphi(X))^{-1}\nabla\phi. \quad (9)$$

Finally, the *Cauchy trajectory* for the function $\phi_{|_{\mathcal{F}^0(\mathrm{P})}}$, with respect to φ, is the differentiable curve $Z : [0,\beta) \to \mathcal{F}^0(\mathrm{P})$ given by

$$Z'(t) = -\operatorname{grad}\phi_{|_{\mathcal{F}^0(\mathrm{P})}}(Z(t)), \qquad Z(0) = Z_0, \quad (10)$$

for the starting point Z_0 and some $\beta > 0$.

Remark 5.1 It is well known that for each $Z_0 \in \mathcal{F}^0(\mathrm{P})$, there exists $\beta > 0$ such that (10) has a unique solution $Z(t)$ defined in $[0,\beta)$.

Consider the following parametrization of the central path $\{X(t) : t \geq 0\}$, where

$$X(t) = \operatorname*{argmin}_{X \succ 0}\{t\langle C, X\rangle + \varphi(X) : \mathcal{A}X = b\}. \quad (11)$$

The next result extends to semidefinite programming, namely, the corresponding one in Iusem et al. [17] obtained for linear programming.

Theorem 5.1 *Let $\{X(t) : t \geq 0\}$ be the central path with respect to φ, as defined in* (11), *and let $\phi(X) = \langle C, X\rangle$. If $Z_0 \in \mathcal{F}^0(\mathrm{P})$ satisfies $\nabla\varphi(Z_0) = \mathcal{A}^* z_0$, for some $z_0 \in \mathbb{R}^m$, then the central path $\{X(t) : t \geq 0\}$ is a solution for* (10) *with $X(0) = Z_0$, i.e., the central path is the Cauchy trajectory for $\phi_{|_{\mathcal{F}^0(\mathrm{P})}}$ in $\mathcal{F}^0(\mathrm{P})$ with respect to φ and starting point Z_0.*

Proof First, note that from optimality condition for (11) we have $tC + \nabla\varphi(X(t)) = \mathcal{A}^* y(t)$, for all $t \geq 0$ and some $y(t) \in \mathbb{R}^m$. So,

$$\nabla\varphi(X(0)) = \mathcal{A}^* y(0).$$

Since $\nabla\varphi(Z_0) = \mathcal{A}^* z_0$ and φ is strictly convex, we have $X(0) = Z_0$ and $y(0) = z_0$. Now, taking derivative in the above equality we obtain

$$C + \nabla^2\varphi(X(t))X'(t) = \mathcal{A}^* y'(t),$$

or equivalently,

$$(\nabla^2\varphi(X(t)))^{-1} C + X'(t) = (\nabla^2\varphi(X(t)))^{-1} \mathcal{A}^* y'(t). \tag{12}$$

Applying $\mathcal{A}$ in this equality, we have

$$\mathcal{A}(\nabla^2\varphi(X(t)))^{-1} C + \mathcal{A}X'(t) = \mathcal{A}(\nabla^2\varphi(X(t)))^{-1} \mathcal{A}^* y'(t).$$

Because $X'(t) \in T_X F^0(\mathrm{P})$, it follows from last equality that

$$\mathcal{A}(\nabla^2\varphi(X(t)))^{-1} C = \mathcal{A}(\nabla^2\varphi(X(t)))^{-1} \mathcal{A}^* y'(t).$$

Now, due the fact that $\mathcal{A}(\nabla^2\varphi(X(t)))^{-1}\mathcal{A}^*$ is nonsingular, is easy to see from latter equality that

$$y'(t) = (\mathcal{A}(\nabla^2\varphi(X(t)))^{-1}\mathcal{A}^*)^{-1}\mathcal{A}(\nabla^2\varphi(X(t)))^{-1} C.$$

Now, substituting $y'(t)$ in (12), we obtain

$$X'(t) = -(I - (\nabla^2\varphi(X(t)))^{-1}\mathcal{A}^*(\mathcal{A}(\nabla^2\varphi(X(t)))^{-1}\mathcal{A}^*)^{-1}\mathcal{A})(\nabla^2\varphi(X(t)))^{-1} C.$$

Finally, as $\phi(X) = \langle C, X\rangle$ it follows from last equation and (9) that $X(t)$ satisfies (10) and the statement of the theorem is proved. □

The next result is a consequence of the latter theorem.

Corollary 5.1 *The central path* $\{X(t) : t \geq 0\}$ *for the problem* (P) *with starting point* $X_0 \in \mathcal{F}^0(\mathrm{P})$, *where* X_0 *satisfies* $\nabla\varphi(X_0) = \mathcal{A}^* y_0$ *for some* $y_0 \in \mathbb{R}^m$, *is bounded.*

Proof Let $\phi(X) = \langle C, X\rangle$ and $\psi(t) = \phi(X(t))$. It follows from Theorem 5.1 that

$$\begin{aligned}\psi'(t) &= \langle \operatorname{grad} \phi|_{\mathcal{F}^0(\mathrm{P})}(X(t)), X'\rangle_{\nabla^2\varphi} = \langle \operatorname{grad} \phi|_{\mathcal{F}^0(\mathrm{P})}(X(t)), -\operatorname{grad} \phi|_{\mathcal{F}^0(\mathrm{P})}(X(t))\rangle_{\nabla^2\varphi} \\ &= \langle \nabla^2\varphi(X(t)) \operatorname{grad} \phi|_{\mathcal{F}^0(\mathrm{P})}(X(t)), -\operatorname{grad} \phi|_{\mathcal{F}^0(\mathrm{P})}(X(t))\rangle \\ &= -\|(\nabla^2\varphi(X(t)))^{1/2} \operatorname{grad} \phi|_{\mathcal{F}^0(\mathrm{P})}(X(t))\|^2 < 0, \quad \forall t \in (0, +\infty).\end{aligned}$$

Then, ϕ is decreasing along to the central path, which implies that $\{X(t) : t > 0\} \subset \{X \in S^n_+; \phi(X) \leq \phi(X_0)\}$. Since that optimal solutions set of (P) is compact, it follows from convexity of ϕ that $\{X \in S^n_+; \phi(X) \leq \phi(X_0), \mathcal{A}X = b\}$ is also compact. Therefore, $\{X(t) : t > 0\}$ is bounded. □

6 Final Remarks

In this paper we have studied the convergence properties of the central path, for semidefinite programming problems, associated to a function satisfying some specific assumptions. We have shown that the central path is well defined and bounded. Moreover, when that function can be continuously extended to the boundary of its domain, we have proved that the central path converges to the analytic center of the solution set of the problem. For a more general class of functions, including the functions presented in the Examples 3.3 and 3.4, the convergence and characterization of the limit point of central path associated to them is an open problem.

As application of the study of the central path, we have presented some convergence properties of the generalized proximal-point method for semidefinite programming problems. Also, convergence results for the generalized proximal-point methods associated to a class of functions including the functions presented in the Examples 3.3 and 3.4 is lacking.

References

1. Alizadeh, F.: Interior point methods in semidefinite programming with applications to combinatorial optimization. SIAM J. Optim. **5**(1), 13–51 (1995)
2. Nesterov, Yu., Nemirovski, A.: Interior Point Polynomial Methods in Convex Programming. SIAM, Philadelphia (1994)
3. Todd, M.J.: Semidefinite optimization. Acta Numer. **10**, 515–560 (2001)
4. Halická, M., de Klerk, E., Roos, C.: On the convergence of the central path in semidefinite optimization. SIAM J. Optim. **12**(4), 1090–1099 (2002)
5. Graña Drummond, L.M., Peterzil, H.Y.: The central path in smooth convex semidefinite programs. Optmization **51**(2), 207–233 (2002)
6. Halická, M., de Klerk, E., Roos, C.: Limiting behavior of the central path in semidefinite optimization. Optim. Methods Softw. **20**(1), 99–113 (2005)
7. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**(5), 877–898 (1976)
8. Chen, G., Teboulle, M.: Convergence analysis of a proximal-like minimization algorithm using Bregman functions. SIAM J. Optim. **3**(3), 538–543 (1993)
9. Eckstein, J.: Nonlinear proximal point algorithm using Bregman functions, with applications to convex programming. Math. Oper. Res. **18**(1), 202–226 (1993)
10. Iusem, A.N.: On some properties of generalized proximal point methods for variational inequalities. J. Optim. Theory Appl. **96**(2), 337–362 (1998)
11. Doljansky, M., Teboulle, M.: An interior proximal algorithm and the exponential multiplier method for semidefinite programming. SIAM J. Optim. **9**(1), 1–13 (1998)
12. Rapcsák, T., Thang, T.T.: A class of polynomial variable-metric algorithms for linear optimization. Math. Program. **74**, 319–331 (1996)
13. Rapcsák, T.: Smooth Nonlinear Optimization in R^n. Kluwer Academic, Dordrecht (1997)
14. Ferreira, O.P., Oliveira, P.R.: Proximal point algorithm on Riemannian manifolds. Optimization **51**(2), 257–270 (2002)
15. Nesterov, Y., Todd, M.J.: On the Riemannian geometry defined by self-concordant barriers and interior-point methods. Found. Comput. Math. **2**(4), 333–361 (2002)
16. Alvarez, F., Bolte, J., Brahic, O.: Hessian Riemannian gradient flows in convex programming. SIAM J. Control Optim. **43**(2), 477–501 (2004)
17. Iusem, A.N., Svaiter, B.F., da Cruz Neto, J.X.: Central paths, generalized proximal point methods and Cauchy trajectories in Riemannian manifolds. SIAM J. Control Optim. **37**(2), 566–588 (1999)
18. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)