CENTRAL PATHS IN SEMIDEFINITE PROGRAMMING AND CAUCHY TRAJECTORIES IN RIEMANNIAN MANIFOLDS

João Xavier da Cruz Neto  
Universidade Federal do Piauí  
Campus Ministro Petrônio Portela, CCN, CEP: 64049-550 - Teresina – Piauí  
jaxvier@ufpi.br

Orizon Pereira Ferreira  
Universidade Federal de Goiás  
Campus Samambaia, IME, CEP: 74001-970 - Goiânia - Goiás  
orizon@mat.ufg.br

Paulo Roberto Oliveira  
Universidade Federal do Rio de Janeiro  
Cidade Universitária, Centro de Tecnologia, PESC, CEP: 21941-972 - Rio de Janeiro – Rio de Janeiro  
poliveir@cos.ufrj.br

Roberto Cristóvão Mesquita Silva  
Universidade Federal do Amazonas  
Campus Universitário, ICE, CEP: 69077-000 - Manaus - Amazonas  
rmesquita@ufam.edu.br

RESUMO

A relação entre a trajetória central no contexto da programação semidefinida e a trajetória de Cauchy na variedade Riemanniana é estudada neste artigo. Primeiramente, prova-se que a trajetória central associada a uma função geral está bem definida. A convergência e a caracterização de seu ponto do limite são estabelecidas para as funções que satisfazem a uma determinada propriedade de continuidade. Finalmente, prova-se que a trajetória central coincide com a trajetória de Cauchy na variedade Riemanniana.

PALAVRAS CHAVE. trajetória central, trajetória de Cauchy, programação semidefinida, programação matemática.

ABSTRACT

The relationship among central path in the context of semidefinite programming and Cauchy trajectory in Riemannian manifolds is studied in this paper. First it is proved that the central path associated to the general function is well defined. The convergence and characterization of its limit point is established for functions satisfying a certain continuous property. Finally, it is proved that the central path coincides with the Cauchy trajectory in the Riemannian manifold.

KEYWORDS. central path, Cauchy trajectory, semidefinite programming, mathematical programming.
1. Introdução

The extension of concepts and techniques from linear programming to semidefinite programming became attractive after the seminal works due to Alizadeh (1995) and Nesterov and Nemirovski (1994). 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 Güler (1994) and Todd (2001). The central path for semidefinite programming problems converges, see Halická et al. (2002). More general, Graña Drummond and Peterzil (2002) 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. (2002). Partial characterizations of the limit point has been given by Luo et al. (1998), Sporre and Forsgren (2002) and Halická et al. (2005).

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 Karmarkar (1990), Smith (1994), Udriste (1994), Rapcsák, T. and Thang (1996), Rapcsák (1997), da Cruz Neto et al. (1998), Ferreira and Oliveira (1998), Ferreira and Oliveira (2002), Nesterov and Todd (2002) and Nesterov and Nemirovski (2003). A couple of paper have dealt with the behavior of the Cauchy trajectories in Riemannian manifolds including Karmarkar (1990), Helmke and Moore (1994), Balogh et al. (2004) and Alvarez et al. (2004).

The central path with respect to general barrier function, for monotone variational inequality problem, has been considered by Iusem et al. (1999) and its well definition and convergence properties was obtained. Characterizations of the limit point for some specific problems including linear programming was given, i.e., it was proved that the central path converges to the analytic center of the solution set. Also, Iusem et al. (1999) provided a connection among central path and Cauchy trajectory (or gradient trajectories) in Riemannian manifold. It was showed that in some cases, including linear programming, these two concepts are in a certain way equivalent.

In this paper we will prove the equivalence among two concepts, namely, central path 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. (1999). 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 continuous 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 Cauchy trajectory in Riemannian manifold.

The organization of our paper is as follows. In Subsection 1.1, we list some basic notation and terminology used in our presentation. In Section 2, we describe the semidefinite programming problem and the basic assumptions that will be used throughout the paper. In Section 3, we introduce some assumptions in order to guarantee well definedness of the central path and establish some results about it. In Section 4, 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 (1985). $\mathbb{R}^n$ denotes the n-dimensional Euclidean space. $\mathbb{R}_+^n = \{ (x_1, \ldots, x_n) \in \mathbb{R}^n ; x_i \geq 0 \ \forall \ i = 1, \ldots, n \}$ and $\mathbb{R}_{++}^n = \{ (x_1, \ldots, x_n) \in \mathbb{R}^n ; x_i > 0 \ \forall \ i = 1, \ldots, n \}$ denote nonnegative and positive orthant,
respectively. The set of all matrices \( n \times n \) is denoted by \( \mathbb{R}^{n \times n} \). The \((i,j)\) -th entry of a matrix \( X \in \mathbb{R}^{n \times m} \) is denoted by \( X_{ij} \) and the \( j \) -th 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 \( \mathbb{S}^n \). The cone of positive semidefinite (resp. definite) \( n \times n \) symmetric matrices is denoted by \( \mathbb{S}^n_+ \) (resp., \( \mathbb{S}^n_+ \)) and \( \partial \mathbb{S}^n_+ \) denotes the boundary of \( \mathbb{S}^n_+ \). The trace of a matrix \( X \in \mathbb{R}^{n \times m} \) is denoted \( \text{tr} \). The Frobenius norm of the matrix \( X \) is defined as \( \|X\| = \sqrt{\text{tr}(X^TX)} \). The image (or range) space and the null space of a linear operator \( A \) will be denoted by \( \text{Im}(A) \) and \( \text{Null}(A) \), respectively; the dimension of the subspace \( \text{Im}(A) \), referred to as the rank of \( A \), will be denoted by \( \text{rank}(A) \). Given a linear operator \( F : E \rightarrow F' \) 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 \( A^* : F' \rightarrow E \) satisfying \( \langle A(u), v \rangle_E = \langle u, A^*(v) \rangle_{F'} \) for all \( u \in E \) and \( v \in F' \). Given \( A, K, A_m \in \mathbb{S}^n \), define the linear application \( A : \mathbb{S}^n \rightarrow \mathbb{R}^m \) by
\[
AX = \left[ \langle A_1, X \rangle, K, \langle A_m, X \rangle \right]^T.
\]Note that the adjoint application \( A^* : \mathbb{R}^m \rightarrow \mathbb{S}^n \) of \( A \) is given by \( A^*v = \sum_{i=1}^{m} v_i A_i \).

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

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

### 2. Preliminares

In this section, we describe our problem and two assumptions on it that will be used throughout the paper. We consider the semidefinite programming problem (SDP)
\[
(P) \quad \min \{ \langle C, X \rangle : AX = b, X \succeq 0 \}
\]
and its associated dual
\[
(D) \quad \max \{ b^T y : A^* y + S = C, S \succeq 0 \},
\]
where the data consist of \( C \in \mathbb{S}^n \), \( b \in \mathbb{R}^m \) and a linear operator \( A : \mathbb{S}^n \rightarrow \mathbb{R}^m \), the primal variable is \( X \in \mathbb{S}^n \), and the dual variable consists of \( (S, y) \in \mathbb{S}^n \times \mathbb{R}^m \). We write \( F(P) \) and \( F(D) \) for the sets of feasible solutions to \( (P) \) and \( (D) \) respectively, and by \( F^0(P) \) and \( F^0(D) \) its relative interior. We also write \( F^*(P) \) and \( F^*(D) \) 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** \( A : \mathbb{S}^n \rightarrow \mathbb{R}^m \) is a surjective linear operator;

**A2** \( F^0(P) \neq \emptyset \) and \( F^0(D) \neq \emptyset \).

Assumption **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. Assumption **A2** ensures that both \( (P) \) and \( (D) \) have optimal solutions, the optimal values of \( (P) \) and \( (D) \) are equals and the solutions
sets $F'(P)$ and $F'(D)$ are limited (see for example Todd (2001)). 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 a function $\varphi : S^n_{++} \to \mathbb{R}^m$. By assuming that $\varphi$ satisfies some assumptions, we prove that the central path is well defined, bounded and converges. Moreover, if $\varphi$ can be continuously extended to $S^n_+$ we prove also that the central path converges to the analytic center of the solution set of the problem $(P)$.

Let $\varphi : S^n_{++} \to \mathbb{R}^m$ be a strictly convex function and $C^2$. We assume this assumption in all our results without mention it. The central path to the Problem $(P)$ with respect to $\varphi$ is the set of points $\{X(\mu) : \mu > 0\}$ defined by

$$X(\mu) = \arg \min_{X \in \mathbb{R}^n} \{(C, X) + \mu \varphi(X) : AX = b\}, \quad \mu \in R_{++}.$$  

Some of our results require one of the following assumptions on $\varphi$.

A3) i) The function $\varphi$ can be continuously extended to $S^n_+$ and for all $\alpha \in R$ the sub-level set

$$L_\alpha = \{X \in S^n_+ : \varphi(X) \leq \alpha\},$$

is bonded;

ii) For all sequence $\{X_k\} \subset S^n_{++}$ such that $\lim_{k \to \infty} X_k = X \in \partial S^n_+$, there holds

$$\lim_{k \to \infty} \langle \nabla \varphi(X_k), \overline{X} - X_k \rangle = -\infty,$$

for all $\overline{X} \in S^n_+$.

A4) i) The function $\varphi$ goes to $+\infty$ as $X$ goes to the boundary $\partial S^n_+$ to $S^n_+$, 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

$$\lim_{\mu \to +\infty} \phi_{V, \mu}(X) = +\infty.$$

The assumptions A3 and A4 are important to assure the well definition of the central path. Now, we are going to show that the assumption A4 implies that $\phi_{V, \mu}$ has a compact sub-level.

Lemma 3.1. Under assumption A4 the sub-level set

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

is compact for each $\alpha \in R$. As consequence, $\phi_{V, \mu}$ has a minimizer in $S^n_+$.

Proof. Let $\alpha \in R$. We claim that $K_{\alpha, \mu}(V)$ is bounded. Indeed, assume by contradiction that $\lim_{X \to \partial S^n_+} \varphi(X) = +\infty$. But, the assumption A4.ii implies $\lim_{X \to \partial S^n_+} \phi_{V, \mu}(X) = +\infty$ which is an absurd, since $\phi_{V, \mu}(X) \leq \alpha$ for all $\mu > 0$. Therefore, $K_{\alpha, \mu}(V)$ is bounded and closed. Since $\phi_{V, \mu}(X)$ is continuous, $K_{\alpha, \mu}(V)$ is compact.
\{X_k\} \subset S^n_{++}, we have two possibilities: \( \overline{X} \in S^n_{++} \) or \( \overline{X} \in \partial S^n_{++} \), where \( \partial S^n_{++} \) denotes the boundary of \( S^n_{++} \). Since \( \mu > 0 \), the assumption \( A4.i \) implies that \( \overline{X} \notin \partial S^n_{++} \). So, \( \overline{X} \in S^n_{++} \).

Now, the continuity of \( \phi_{\nu,\mu} \) in \( S^n_{++} \) implies that \( \alpha \geq \lim_{k \to +\infty} \phi_{\nu,\mu}(X_k) = \phi_{\nu,\mu}(\overline{X}) \) i.e., \( \overline{X} \in K_{\alpha,\mu} \). Thus, \( K_{\alpha,\mu}(V) \) is closed. Therefore \( K_{\alpha,\mu}(V) \) is compact, and it easy to conclude that \( \phi_{\nu,\mu} \) has a minimizer in \( S^n_{++} \).

The above assumptions will be applied to the function \( \varphi : S^n_{++} \to R \) in the following examples:

**Example 3.1.** Let \( \varphi : S^n_{++} \to R \) be given by \( \varphi(X) = \text{tr}(X \ln(X)) \). Clearly \( \varphi \) extends continuously to \( S^n_{++} \) with the convention that \( 0 \ln 0 = 0 \). The gradient of \( \varphi \) is given by \( \nabla \varphi(X) = \text{tr} \left( X^{-1} \right) \). It is easy to see that the function \( \varphi \) is strictly convex and has a unique minimizer \( X^* = e^{-I} \). Therefore, \( L_{\alpha} = \{X \in S^n_{++} : \varphi(X) \leq \alpha \} \) is bounded. Let us consider \( \{X_k\} \subset S^n_{++} \) such that \( \lim_{k \to +\infty} X_k = X \in \partial S^n_{++} \) and \( \overline{X} \in S^n_{++} \). So,

\[
\phi_{\nu,\mu}(X_k) = \left( \ln(X_k) + I, \overline{X} - X_k \right) = \left( \ln(X_k), \overline{X} \right) - \left( \ln(X_k), X_k \right) + \left( I, \overline{X} \right) - \left( I, X_k \right) \leq \sum_{i=1}^{n} \lambda_i(\overline{X})\lambda_i(\ln(X_k)) - \varphi(X_k) + \sum_{i=1}^{n} \lambda_i(\overline{X}) - \sum_{i=1}^{n} \lambda_i(X_k).
\]

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

For details on properties of \( \ln(X) \) see Horn and Johnson (1991).

**Example 3.2.** Let \( \varphi : S^n_{++} \to R \) be given by \( \varphi(X) = -\ln \det(X) \). It easy to see that \( \varphi \) is strictly convex. So, for each \( V \in S^n_{++} \) and \( \mu > 0 \) the function \( \phi_{\nu,\mu}(X) = \langle V, X \rangle - \mu \ln \det(X) \) is also strictly convex. Since \( \phi_{\nu,\mu} \) is convex we obtain \( \phi_{\nu,\mu}(X) \geq \phi_{\nu,\mu}(2\mu V^{-1}) + \langle \nabla \phi_{\nu,\mu}(2\mu V^{-1}), X - 2\mu V^{-1} \rangle \) or equivalently

\[
\phi_{\nu,\mu}(X) \geq \frac{1}{2} \langle V, X \rangle + \mu n - \mu \ln \det(2\mu V^{-1}).
\]

As \( V \in S^n_{++} \) we have \( \lim_{x \to \partial S^n_{++}} \langle V, X \rangle = +\infty \) thus latter inequality implies \( \lim_{x \to \partial S^n_{++}} \phi_{\nu,\mu}(X) = +\infty \). Now,

\[
\lim_{X \to \partial S^n_{++}} - \ln \det(X) = +\infty.
\]

Therefore, \( \varphi \) satisfies \( A4 \).

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

\[
\nabla \varphi(X) = -\alpha \det X^{-\alpha} X^{-1}
\]

and

\[
\nabla^2 \varphi(X) H = \alpha \det X^{-\alpha} \left( X^{-1} H X^{-1} + X^{-1} H X^{-1} \right)
\]

where \( H \in S^n \). Hence we obtain
\[ \langle \nabla^2 \varphi(X) H, H \rangle = \alpha \det X^{-\alpha} \left( \alpha \langle X^{-1}, H \rangle^2 + \left\| X^{-\frac{\alpha}{2}} H X^{-\frac{\alpha}{2}} \right\|^2 \right) > 0, \]

for all \( H \neq 0 \). So, \( \varphi \) is strictly convex, as a consequence \( \phi_{\nu, \mu}(X) = \langle V, X \rangle + \mu \det X^{-\alpha} \) is also strictly convex, for all \( V \in S^a_+ \) and \( \mu > 0 \). Now, since \( \mu > 0 \), \( \det X^{-\alpha} > 0 \) and \( V \in S^a_+ \) we have

\[
\lim_{H \to \infty} \phi_{\nu, \mu}(X) \geq \lim_{H \to \infty} \langle V, X \rangle = +\infty.
\]

Finally, as \( \lim_{X \to \partial S^+_n} \det X^{-\alpha} = +\infty \), we have that \( \varphi \) satisfies A4.

**Example 3.4.** Let \( \varphi : S^a_+ \to R \) be given by \( \varphi(X) = \text{tr } X^{-1} \). So, \( \nabla \varphi(X) = -X^{-2} \) and \( \nabla^2 \varphi(X) H = X^{-2} H X^{-1} + X^{-1} H X^{-2} \), where \( H \in S^a \). Then, we obtain

\[
\langle \nabla^2 \varphi(X) H, H \rangle = \left\| X^{-1} H X^{-\frac{1}{2}} \right\|^2 + \left\| X^{-\frac{1}{2}} H X^{-1} \right\|^2 > 0,
\]

for all \( H \neq 0 \), which implies that \( \varphi \) is strictly convex. Thus, \( \phi_{\nu, \mu}(X) = \langle V, X \rangle + \mu \text{ tr } X^{-1} \) is also strictly convex, for all \( V \in S^a_+ \) and \( \mu > 0 \). As, \( \text{tr } X^{-1} > 0 \), \( \mu > 0 \) and \( V \in S^a_+ \) we have

\[
\lim_{\mu \to \infty} \phi_{\nu, \mu}(X) \geq \lim_{\mu \to \infty} \langle V, X \rangle = +\infty.
\]

Now, since \( \lim_{X \to \partial S^+_n} \text{tr } X^{-1} = +\infty \), we obtain that \( \varphi \) satisfies A4.

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

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

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

First, assume that \( \varphi \) satisfies A3. It easy to see that \( \langle C, X \rangle = \langle S_0, X \rangle + b^T y_0 \), for all \( X \in F(P) \), hence (2) is equivalent to

\[
X(\mu) = \arg \min_{X \in F(0)} \{ \phi_{(S_0, \mu)}(X) : AX = b, \phi_{(S_0, \mu)}(X) \leq \phi_{(S_0, \mu)}(X_0) \}.
\]

Let us consider the sub-level set \( L_\beta = \{X \in S^a_+ : \phi_{(S_0, \mu)}(X) \leq \beta\} \), where \( \beta = \phi_{(S_0, \mu)}(X_0) \). Note that \( L_\beta \subset L_\alpha = \{X \in S^a_+ : \varphi(X) \leq \alpha\} \), where \( \alpha = \beta / \mu \). From A3.i we have that \( L_\alpha \) is bounded. So, \( L_\beta \) is also bounded. As \( \phi_{(S_0, \mu)} \) is continuous in \( S^a_+ \) we have that \( L_\beta \) is compact, which implies that \( L_\beta \) is also bounded. As \( \phi_{(S_0, \mu)} \) is strictly convex in \( S^a_+ \) we have that \( L_\beta \) is compact. Since \( \phi_{(S_0, \mu)} \) is strictly convex we have that there exists a unique minimizer \( X(\mu) \in F(P) \) 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 \( X(\mu) \in F(0)(P) \). Assume by contradiction that \( X(\mu) \in \partial F(P) = \{X \in \partial S^a_+ : AX = b\} \) and 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 \( X_0 \in F(0) \), \( X(\mu) \in \partial F(P), \xi_k \in (0,1) \) and \( F(0)(P) \) is convex, we conclude that \( Z_k \in F(0)(P) \) for all \( \varepsilon_k \in (0,1) \). Now, combining definitions of \( X(\mu) \) and sequence \( \{Z_k\} \) with convexity of \( \varphi \) we obtain
\[
0 \leq \langle C, Z_k \rangle + \mu \varphi(Z_k) - \langle C, X(\mu) \rangle - \mu \varphi(X(\mu)) \\
\leq \langle C, Z_k - X(\mu) \rangle + \mu \langle \nabla \varphi(Z_k), Z_k - X(\mu) \rangle \\
= \varepsilon_k \langle C, X_0 - X(\mu) \rangle + \mu \frac{\varepsilon_k}{1 - \varepsilon_k} \langle \nabla \varphi(Z_k), X_0 - Z_k \rangle.
\]

The latter inequality implies that
\[
\frac{(1 - \varepsilon_k)}{\mu} \langle C, X(\mu) - X_0 \rangle \leq \langle \nabla \varphi(Z_k), X_0 - Z_k \rangle.
\]

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

Finally, assume that \(\varphi\) satisfies A4. Let us consider the sub-level set
\[
K_{\alpha, \mu}(S_0) = \{X \in S^n_+: \phi_{S_0, \mu}(X) \leq \alpha \},
\]
where \(\alpha = \phi_{(S_0, \mu)}(X_0)\). Now, from Lemma 3.1 it easy to see that \(K_{\alpha, \mu}(S_0) \cap \{X \in S^n_+: AX = b\}\) is compact. So, a similar argument used in the first part permit to conclude that the central path \(\{X(\mu) : \mu > 0\}\) is well defined and definition of \(K_{\alpha, \mu}(S_0)\) implies that it is in \(F^0(P)\).

For \(\varphi\) satisfying A3 or A4, the above Theorem implies that the central path \(\{X(\mu) : \mu > 0\}\), with respect to \(\varphi\), is well defined and is in \(F^0(P)\). So, for all \(\mu > 0\), we have from (2) that
\[
\mu \nabla \varphi(X(\mu)) = -C + A^* y(\mu),
\]
for some \(y(\mu) \in \mathbb{R}^m\).

**Proposition 3.1.** Suppose that \(\varphi\) satisfies A3 or A4. Then, the following statements hold:

i) the function \(0 < \mu \alpha \varphi(X(\mu))\) is non-increasing;

ii) the set \(\{X(\mu) : 0 < \mu < \mu_0\}\) is bounded, for each \(\mu_0 > 0\);

iii) all cluster points of \(\{X(\mu) : \mu > 0\}\) are solutions of the Problem \((P)\).

**Proof.** The proof of the items i), ii) and iii) are similar to the proof of the Proposition 3.i, Proposition 4 and Proposition 5 of Iusem et al. (1999), respectively.

**Theorem 3.2.** Assume that \(\varphi\) satisfies A3. Let \(X^* \in S^n_+\) be the analytic center of \(F^*(P)\) with respect to \(\varphi\), i.e., the unique solution of the problem
\[
(S) \begin{cases}
\min_X \varphi(X) \\
\text{s.t. } X \in F^*(P).
\end{cases}
\]
Then \(\lim_{\mu \to 0} X(\mu) = X^*\), where \(\{X(\mu) : \mu > 0\}\) is the central path with respect to \(\varphi\).

**Proof.** Take \(X^*\) 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 \(\lim_{k \to \infty} X(\mu_k) = X^*\). Now, from (4) we have
\[
C + \mu_k \nabla \varphi(X(\mu_k)) = A^* y(\mu_k),
\]
for some \(y(\mu_k) \in \mathbb{R}^m\). So,
\[
\langle \mu_k \nabla \varphi(X(\mu_k)), X^* - X(\mu_k) \rangle = \langle A^* y(\mu_k), X^* - X(\mu_k) \rangle,
\]
for all \(X \in F^*(P)\). Using that \(X - X(\mu_k) \in \text{Null}(A)\), this equality becomes
\[ \left\langle \mu_k \nabla \phi(X(\mu_k)), X - X(\mu_k) \right\rangle = -\left\langle C, X - X(\mu_k) \right\rangle. \]

Since \( \phi \) is convex, the above equality implies that
\[ \mu_k (\phi(X(\mu_k)) - \phi(X)) \leq \left\langle C, X - X(\mu_k) \right\rangle. \]

Because \( X \in F^*(P) \) and \( \mu_k > 0 \), it follows from the latter inequality that \( \phi(X(\mu_k)) \leq \phi(X) \).

Now, as \( \phi \) is continuous we can take limits, as \( k \) goes to \( +\infty \), in this inequality to conclude that \( \phi(\hat{X}) \leq \phi(X) \), for all \( X \in F^*(P) \). Thus, any cluster point \( \hat{X} \) of \( \{X(\mu) : \mu > 0\} \) is a solution of the problem (S).

Therefore, since \( X^* \) is the unique solution of the problem (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 \( \phi(X) = -\ln \det(X) \) converges, as \( \mu \) goes to 0.


**4. 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 \( \phi \) for the Problem (P), becomes a Cauchy trajectory on the Riemannian manifold endowed with the metric given by the Hessian of \( \phi \). This result extends to semidefinite programming context the corresponding result of linear programming, see Section 4 of Iusem et al. (1999).

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 \( \left\langle \cdot, \cdot \right\rangle \). The tangent space to \( S^n_+ \) at \( X \) is given by \( T_X S^n_+ = \{Y - X; Y \in S^n\} = S^n_+ \). Let \( \phi : S^n_+ \to R \) be strictly convex and \( C^2 \). Define a new metric in \( S^n_+ \) as
\[ \left\langle U, V \right\rangle_{\nabla^2 \phi(X)} = \left\langle \nabla^2 \phi(X) U, V \right\rangle, \]
so that \( M := (S^n_+, \nabla^2 \phi) \) is now a Riemannian Manifold. The metric of \( M \) induces a map \( \phi \alpha \) \( \text{grad} \phi \) which associates to each \( \phi \in C^1(S^n_+) \) its gradient \( \text{grad} \phi \in S^n \) by the rule
\[ d\phi_X(V) = \left\langle \text{grad} \phi(X), V \right\rangle_{\nabla^2 \phi(X)}, \]
where \( X \in S^n_+ \) and \( V \in S^n \). It is easy to see that the gradient vector field in \( M \) is given by
\[ \text{grad} \phi(X) = (\nabla^2 \phi(X))^{-1} \nabla \phi(X), \]
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 \( A : S^n \to R^n \) be the linear operator as defined in (1). The assumptions A1 and A2 imply that the set \( F^0(P) = \{X \in S^n : AX = b, X \in S^n_+\} \) is a Riemannian submanifold of \( M \) with the induced metric and tangent space at \( X \) given by \( T_X F^0(P) = \{X \in S^n : AV = 0\} \). The adjoint operator of \( A \) with respect to the metric of \( M \) is \( (\nabla^2 \phi(X))^{-1} A^* \), where \( A^* : R^n \to S^n \) is the usual adjoint operator of \( A \). In this case, the orthogonal projection \( \Pi_X : S^n \to T_X F^0(P) \) with respect to the metric of \( M \) is
\[ \Pi_X = I - (\nabla^2 \phi(X))^{-1} A^*(A(\nabla^2 \phi(X))^{-1} A^*)^{-1} A. \]

The gradient vector field of the function \( \phi |_{F^0(P)} : F^0(P) \to R \), with respect to the metric of \( M \),
is given by
\[ \nabla \phi |_{\nu} = \Pi \nabla \phi, \text{ i.e.,} \]
\[ \nabla \phi |_{\nu} = (I - (V^2 \phi)^{-1} A^* (A (V^2 \phi)^{-1} A^* A^*)^{-1} A^* A^*)^{-1} V^2 \phi. \tag{5} \]

Finally, the Cauchy trajectory for the function \( \phi|_{v} \), with respect to \( \phi \), is the differentiable curve \( Z: [0, \beta) \to F^0(P) \) given by
\[ \begin{align*}
Z(t) &= -\nabla \phi |_{\nu} (Z(t)) \\
Z(0) &= Z_0.
\end{align*} \tag{6} \]

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

**Remark 4.1.** It is well known that for each \( Z_0 \in F^0(P) \), there exists \( \beta > 0 \) such that (6) has a unique solution \( Z(t) \) defined in \( [0, \beta) \).

Consider the following parametrization of the central path \( \{X(t): t \geq 0\} \) where
\[ X(t) = \arg \min_{\nu \in F^0(P)} \{ C, X \} \phi(X) : A^* y = b \}. \tag{7} \]

Next result extends to semidefinite programming, namely, the corresponding one in Iusem et al. (1999) obtained for linear programming.

**Theorem 4.1.** Let \( \{X(t): t \geq 0\} \) be the central path with respect to \( \phi \), as defined in (7), and let \( \phi(X) = \langle C, X \rangle \). If \( Z_0 \in F^0(P) \) satisfies \( \nabla \phi(Z_0) = A^* z_0 \), for some \( z_0 \in \mathbb{R}^m \), then the central path \( \{X(t): t \geq 0\} \) is a solution for (6) with \( X(0) = Z_0 \), i.e, the central path is the Cauchy trajectory for \( \phi|_{\nu} \) in \( F^0(P) \) with respect to \( \phi \) and starting point \( Z_0 \).

Proof. First, note that from optimality condition for (7) we have \( IC + \nabla \phi(X(t)) = A^* y(t) \), for all \( t \geq 0 \) and some \( y(t) \in \mathbb{R}^m \). So, \( \nabla \phi(X(0)) = A^* y(0) \). Since \( \nabla \phi(Z_0) = A^* z_0 \) and \( \phi \) is strictly convex we have \( X(0) = Z_0 \) and \( y(0) = z_0 \). Now, taking derivative in the above equality we obtain \( C + \nabla^2 \phi(X(t)) X'(t) = A^* y'(t) \), or equivalently
\[ (\nabla^2 \phi(X(t)))^{-1} C + X'(t) = (\nabla^2 \phi(X(t)))^{-1} A^* y'(t). \tag{8} \]

Applying \( A \) in this equality we have
\[ A(\nabla^2 \phi(X(t)))^{-1} C + AX'(t) = A(\nabla^2 \phi(X(t)))^{-1} A^* y'(t). \]

Because \( X'(t) \in T_X F^0(P) \), it follows from last equality that
\[ A(\nabla^2 \phi(X(t)))^{-1} C = A(\nabla^2 \phi(X(t)))^{-1} A^* y'(t). \]

Now, due the fact that \( A(\nabla^2 \phi(X(t)))^{-1} A^* \) is nonsingular, is easy to see from latter equality that
\[ y'(t) = \left(A(\nabla^2 \phi(X(t)))^{-1} A^* \right)^{-1} A(\nabla^2 \phi(X(t)))^{-1} C. \]

Now, substituting \( y'(t) \) in (8) we obtain
\[ X'(t) = -\left(I - (\nabla^2 \phi(X(t)))^{-1} A^* (A(\nabla^2 \phi(X(t)))^{-1} A^* A^*)^{-1} A^* \right)(\nabla^2 \phi(X(t)))^{-1} C. \]

Finally, as \( \phi(X) = \langle C, X \rangle \) it follows from last equation and (5) that \( X(t) \) satisfies (6) and the statement of the theorem is proved.
The next result is a consequence of the latter theorem.

**Corollary 4.1.** The central path \( \{X(t) : t \geq 0\} \) for the problem \( (P) \) with starting point \( X_0 \in P^0(P) \) where \( X_0 \) satisfies \( \nabla \phi(X_0) = A^*y_0 \) for some \( y_0 \in R^m \), is bounded.

**Proof.** Let \( \phi(X) = \langle C, X \rangle \) and define \( \psi(t) = \phi(X(t)) \). It follows from Theorem 4.1 that

\[
\psi'(t) = \left( \text{grad } \phi \mid _{\psi(t)} (X(t)) , X'(t) \right)_{\nabla^2 \phi} = \left( \text{grad } \phi \mid _{\psi(t)} (X(t)) , -\text{grad } \phi \mid _{\psi(t)} (X(t)) \right)_{\nabla^2 \phi} = \left( \nabla^2 \phi(X(t)) \text{grad } \phi \mid _{\psi(t)} (X(t)) , -\text{grad } \phi \mid _{\psi(t)} (X(t)) \right) = -\left\| \nabla^2 \phi(X(t)) \right\|^{1/2}_{\nabla^2 \phi} \text{grad } \phi \mid _{\psi(t)} (X(t)) \left\| < 0, \forall t \in (0, +\infty) \right. .
\]

Then \( \phi \) 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 \( \phi \) that \( \{X \in S^n_+ : \phi(X) \leq \phi(X_0), AX = b\} \) is also compact. Therefore, \( \{X(t) : t > 0\} \) is bounded.

4. Final Remarks

In this paper we have studied the convergence properties of central path, for semidefinite programming problems, associated to a function satisfying some specific assumptions. We have showed that the central path is well defined and bounded. Moreover, whether 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. Partial characterizations of the limit point of the central path with respect to the log-barrier function for semidefinite programming problems has been obtained by Sporre and Forsgren (2002), Halická et al. (2005) and da Cruz Neto et al. (2005). 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 that it coincides with the Cauchy trajectory in Riemannian Manifolds.

**References**