A Sequential Quadratic Penalty Method for Nonlinear Semidefinite Programming[1]

X. X. Huang[2]

Department of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047, China

X. Q. Yang[3] **and K. L. Teo**

Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong

Abstract. In this paper, a nonlinear semidefinite program is reformulated into a mathematical program with a matrix equality constraint and a sequential quadratic penalty method is proposed to solve the latter problem. We discuss the differentiablity and convexity of the penalty function. Necessary and sufficient conditions for the convergence of optimal values of penalty problems to that of the original semidefinite program are obtained. The convergence of optimal solutions of penalty problems to that of the original semidefinite program is also investigated. We show that any limit point of a sequence of stationary points of penalty problems satisfies the KKT optimality condition of the semidefinite program. Smoothed penalty problems that have the same order of smoothness as the original semidefinite program are adopted. Corresponding results such as the convexity of the smoothed penalty function, the convergence of optimal values, optimal solutions and the sationary points of the smoothed penalty problems are obtained.

Key words: Semidefinite program, penalty method, smoothing method, optimality condition, convergence.

[1]This work is supported by the Postdoctoral Fellowship of Hong Kong Polytechnic University.
[2]Current Address: Department of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong.
[3]Corresponding author.

1 Introduction

It is well-known that semidefinite programming has wide applications in engineering, economics and combinatorial optimization and has received considerable attention in the optimization community (see, e.g., [1, 30, 12] and the references therein). Recent research shows that semidefinite programming is also very useful in nonconvex quadratic optimization (see, [19, 29, 35, 36] and the references therein). Linear semidefinite programs are mainly solved by interior-point algorithms (see, e.g., [30, 34, 31, 28, 17] and the references therein). Nonlinear semidefinite programming arises in optimal structural design (see [21, 22]), optimal robust control (see [9, 11]) and feedback control (see [7, 11]). For a comprehensive review of the applications of nonlinear (nonconvex) semidefinite programs, we refere the reader to [2, 15]. In comparison with linear semidefinite programming, the study of nonlinear semidefinite programming, in particular, nonconvex semidefinite programming, is somewhat limited (see [20, 26, 23, 24, 3, 4, 8, 2, 14, 15]). Recently, a class of penalty/barrier multiplier methods was proposed for the solution of convex semidefinite programming with a linear matrix inequality constraint (see [18]). Most recently, a class of linear and nonlinear semidefinite programs are reformulated into nonlinear programs. As a result, this class of semidefinite programs can be solved through the solution of the reformulated nonlinear programs (see [5, 6]). Barrier methods were suggested for the general (SDP) in [21, 22, 2, 14, 15]. These methods require a strict (interior) fesaible solution as the starting point, which is not easy to be found even if it exists.

It is well-known that sequential penalty method is an important method for constrained nonlinear programming (see, e.g., [10]). Compared with barrier methods, penalty methods are more robust and need not start with a feasible point. In this paper, we shall reformulate a general nonlinear semidefinite program into a mathematical program with a nonsmooth matrix equality constraint and then apply a sequential quadratic penalty method to the reformulated problem.

Let S_m be the set of $m \times m$ real symmetric matrices and for $A \in S_m$, the notation $A \succeq 0$ means that A is positive semidefinite. By $A \not\succeq 0$, we mean that A is not positive semidefinite. Let $A, B \in S_m$. We write $A \succeq B$ if and only if $A - B \succeq 0$. Let $A \succeq 0$. Denote by $A^{1/2}$ or $\sqrt{A}$ the unique (positive semidefinite) square root of A. For $A \in S_m$, define $|A| = (A^2)^{1/2}$. If A is nondegenerate, denote by A^{-1} or $1/A$ the inverse of A. Denote $A \succ 0$ if and if A is positive definite.

Consider the following nonlinear semidefinite program:

$$\text{(SDP)} \quad \min f(x)$$
$$\text{s.t. } x \in R^n$$

$$g(x) \succeq 0,$$

where $f : R^n \to R$, $g : R^n \to S_m$ are continuously differntiable.

Suppose that X and Y are two normed spaces. Let $h : X \to Y$ be a (Fréchet) differentiable operator. Let $x \in X$. We use $Dh(x)$ to denote the (Fréchet) derivative of h at x. Let $d \in X$. We use $Dh(x)(d)$ to denote the directional derivative of h at x in the direction d.

Denote by X_0 the feasible set of (SDP), i.e., $X_0 = \{x \in R^n : g(x) \succeq 0\}$. Throughout the paper, we assume that $X_0 \neq \emptyset$.

Note that $A \succeq 0$ if and only if $|A| - A = 0$ ([27]). It follows that (SDP) can be reformulated as the following equivalent constrained optimization problem:

$$\begin{aligned} \text{(P)} \quad & \min f(x) \\ & \text{s.t. } x \in R^n \\ & \qquad |g(x)| - g(x) = 0. \end{aligned}$$

A solution scheme for (P) is to solve the following quadratic penalty problem:

$$(PP_r) \quad \min F(x, r) =: f(x) + r\||g(x)| - g(x)\|^2,$$

where $r > 0$ is the penalty parameter and the norm $\|\cdot\|$ is the Frobenius norm of an $m \times m$ matrix, i.e., $\|A\| = \sqrt{\text{trace}(A^T A)}$, for any $m \times m$ matrix A.

It is clear from [?] that the symmetric-matrix-valued function $|X|X$ is continuously differentiable on S_m. As a result, the real-valued function $\||g(x)| - g(x)\|^2 = 2\text{trace}(g^2(x)) - 2\text{trace}(|g(x)|g(x))$ is also continuously differentiable. However, we note that the term $\||g(x)| - g(x)\|^2$ in the objective function of (PP_r) may not be twice continuously differentiable no matter how highly smooth the symmetric-matrix-valued function $g(x)$ is. This fact prevents the application of the popular Newton method to solve (PP_r) when the data of (SDP) are twice continuously differntiable. On the other hand, the matrix $g(x)$ may be singular, which prevents us from invoking of the function"sqrtm(X)" (to compute $|g(x)| = \sqrt{g^2(x)}$) if we use the MATLAB code to solve (PP_r) directly. These considerations lead us to adopt the following smoothing scheme for (PP_r):

$$(PP_r^{\epsilon_r}) \quad f(x) + r\|\sqrt{g^2(x) + \epsilon_r^2 I} - g(x)\|^2,$$

where $\epsilon_r > 0$ is a scalar satisfying $r\epsilon_r^2 \to 0$ as $r \to +\infty$ and $I \in S_m$ is the identity matrix.

The outline of the paper is as follows. In Section 2, we investiagte the differentiablity and convexity of the objective function of the penalty problem (PP_r) and the convexity of the smoothed penalty problem $(PP_r^{\epsilon_r})$. In Section 3, we study necessary and sufficient conditions for the convergence of optimal values of the penalty problems (PP_r) $((PP_r^{\epsilon_r}))$ to that of (SDP). Some sufficient conditions will also be given to guarantee the existence and convergence of the optimal solutions of the penalty problems (PP_r) $((PP_r^{\epsilon_r}))$. In Section 4, we derive necessary optimality conditions for a local solution of the penalty problem (PP_r) $((PP_r^{\epsilon_r}))$. Section 5 deals with the convergence of stationary points of the penalty problems (PP_r) $((PP_r^{\epsilon_r}))$. Section 6 concludes the paper.

2 Some Basic Properties of Penalty Problems

In this section, we discuss some basic issues such as the differentiablity and the convexity of penalty problems (PP_r) and $(PP_r^{\epsilon_r})$.

Definition 2.1. Let $h : R^n \to S_m$. We say that h is convex on R^n if for any $\theta \in [0,1]$ and any $x_1, x_2 \in R^n$, there holds $h(\theta x_1 + (1-\theta)x_2) \preceq \theta h(x_1) + (1-\theta)h(x_2))$.

It is elementary to verify that h is convex if and only if for any $\Lambda \succeq 0$, the function $\text{trace}(\Lambda h) : R^n \to R$ is convex.

First we deal with the differentiablity of the objective function of (PP_r).

We need the following lemma, which was proved in [25].

Lemma 2.1. Let $f_0 : R \to R$ be continuously differentiable. Define $F : S_m : S_m$ by $F(X) = U^T f_0(\Lambda)U$, where $X = U^T \Lambda U$ is the spectral decomposition of X. Then, F is also continuously differentiable, and for any $Y \in S_m$,

$$DF(X)(Y) = A(X,Y) + \sum_i f_0'(\lambda_i) P_i Y P_i,$$

where

$$A(X,Y) = \frac{1}{2} \sum_{i \neq j} \frac{f_0(\lambda_i) - f_0(\lambda_j)}{\lambda_i - \lambda_j} (P_i Y P_j + P_j Y P_i)$$

and λ_i are different eigenvalues of X and P_i is the projection onto the eigenspace corresponding to λ_i.

Now we have the next result.

Lemma 2.2. Let F be defined as in Lemma 2.1. Define $\psi(X) = \text{trace}(F(X))$. Then, for any $Y \in S_m$, we have

$$D\psi(X)(Y) = \text{trace}(F'(X)Y),$$

where $F'(X) = U^T f_0'(\Lambda)U$.

Proof. Note that trace$(A(X,Y)) = 0$ since $P_iP_j = 0$ if $i \neq j$. The conclusion follows from Lemma 2.1 and this observation. □

The following lemma is a direct consequence of Lemma 2.2.

Lemma 2.3. Let $h_1(X) = \text{trace}\left[(X + |X|)^2\right]$, $h_2(X) = \text{trace}\left[(\sqrt{X^2 + \epsilon^2 I} + X)^2\right]$, $X \in S_m$. Then, both h_1 and h_2 are continuously differentiable (in fact, $C^{1,1}$) on S_m, and for any $Y \in S_m$,

$$Dh_1(X)(Y) = 2\text{trace}[(X + |X|)Y),$$

$$Dh_2(X)(Y) = 2\text{trace}\left[\frac{\left(X + \sqrt{X^2 + \epsilon^2 I}\right)^2}{\sqrt{X^2 + \epsilon^2 I}} Y\right].$$

The following proposition follows from Lemma 2.3 and the chain rule.

Proposition 2.1. Suppose that $g : R^n \to S_m$ is continuously differentiable. Then the function trace$\left((|g(x)| - g(x))^2\right)$ is continuously differentiable on R^n and

$$D\left[\text{trace}(|g| - g)^2\right](x)(d) = 2\text{trace}\left(g(x)Dg(x)(d)\right) - 2\text{trace}\left(|g(x)|Dg(x)(d)\right). \qquad (1)$$

Furthermore, if g is $C^{1,1}$ on R^n, then trace$\left((|g(x)| - g(x))^2\right)$ is also $C^{1,1}$ on R^n.

By Proposition 2.1, it is clear that if the functions involved in (SDP) are continuously differentiable (resp. $C^{1,1}$), then the objective function of penalty problem (PP_r) is also continuously differentiable (resp. $C^{1,1}$).

Now we consider the convexity of the objective function of $(PP_r^{\epsilon_r})$ if f and $-g$ are convex. We need the next lemma, which follows immediately from Theorem 2.3.14 of [16].

Lemma 2.4. Let h_1 and h_2 be defined as in Lemma 2.2. Then, both h_1 and h_2 are convex on S_m.

Now we prove the following lemma.

Lemma 2.5. Let h_1 and h_2 be defined as in Lemma 2.2. Let $X_1 \preceq X_2$. Then, $h_i(X_1) \leq h_i(X_2), i = 1, 2$.

Proof. It is clear from Lemma 2.2 that $Dh_i(X) \succeq 0, i = 1, 2$. Moreover, by the convexity of h_i, we have

$$h_i(X_2) - h_i(X_1) = \text{trace}(Dh_i(X_1)(X_2 - X_1)) \geq 0.$$

The proof is complete. □

The next proposition shows that if (SDP) is a convex programming, then penalty problems (PP_r) and $(PP_r^{\epsilon_r})$ are also convex.

Proposition 2.2. Suppose that f and $-g$ are convex on R^n. Then the objective functions of penalty problem (PP_r) and $(PP_r^{\epsilon_r})$ are also convex.

Proof. We only prove that the objective function of penalty problem $(PP_r^{\epsilon_r})$is convex since the case of (PP_r) can analogously proved.

It is enough to show that $h_2(x) = \|\sqrt{g^2(x) + \epsilon_r^2 I} - g(x)\|^2$ is convex on R^n. Let $\alpha \in [0, 1]$ and $x_1, x_2 \in R^n$. By the convexity of $-g$, we have

$$-g(\alpha x_1 + (1 - \alpha)x_2) \preceq -\alpha g(x_1) - (1 - \alpha)g(x_2).$$

This combined with Lemmas 2.4 and 2.5 yields

$$\begin{aligned} h_2(\alpha x_1 + (1 - \alpha)x_2) &= \|\sqrt{(-g(\alpha x_1 + (1 - \alpha)x_2))^2 + \epsilon_r^2 I} - g(\alpha x_1 + (1 - \alpha)x_2)\|^2 \\ &\leq \|\sqrt{(\alpha g(x_1) + (1 - \alpha)g(x_2))^2 + \epsilon_r^2 I} - (\alpha g(x_1) + (1 - \alpha)g(x_2))\|^2 \\ &\leq \alpha h_2(X_1) + (1 - \alpha)h_2(x_2). \end{aligned}$$

3 Convergence Analysis of Optimal Values and Optimal Solutions

In this section, we give necessary and sufficient conditions that guarantee the convergence of optimal values of (PP_r) $((PP_r^{\epsilon_r}))$ to that of (SDP) as $r \to +\infty$. We also investigate the convergence of optimal solutions of (PP_r) $((PP_r^{\epsilon_r}))$ to that of (SDP) as $r \to +\infty$.

Consider the perturbed problem of (SDP):

$$\begin{aligned} (SDP_u) \quad & \min f(x) \\ & \text{s.t. } x \in R^n \\ & \qquad g(x) + uI \succeq 0, \end{aligned}$$

where $u \geq 0$ is a scalar. Denote by $v(u)$, $v_1(r)$, $v_2(r, \epsilon_r)$ the optimal values of problems (SDP_u), (PP_r) and $(PP_r^{\epsilon_r})$, respectively. Then, it is obvious that $v(0)$ is the optimal value of the problem (SDP).

3.1 Penalty Problems (PP_r)

In this subsection, we discuss the convergence of optimal values and optimal solutions of (PP_r).

Theorem 3.1. Assume that there exists $\bar{r} > 0$ and $m_0 \in R$ such that

$$F(x, \bar{r}) \geq m_0, \quad \forall x \in R^n. \tag{2}$$

Then, $\lim_{r \to +\infty} v_1(r) = v(0)$ if and only if $\liminf_{u \to 0^+} v(u) = v(0)$.

Proof. Sufficiency. Suppose to the contrary that there exist $0 < r_k \to 0$ and $\delta > 0$ such that

$$v_1(r_k) \leq v(0) - \delta, \forall k.$$

It follows that there exists x_k such that

$$\begin{aligned} m_0 + (r_k - \bar{r})\||g(x_k)| - g(x_k)\|^2 &\leq f(x_k) + r_k\||g(x_k)| - g(x_k)\|^2 \\ &\leq v(0) - \delta/2, \quad \forall k. \end{aligned} \tag{3}$$

As a result,

$$\||g(x_k)| - g(x_k)\|^2 \leq \frac{v(0) - m_0 - \delta/4}{r_k - \bar{r}} = \tau_k. \tag{4}$$

Suppose that

$$g(x_k) = U_k^T \text{diag}(\lambda_{1,k}, \cdots, \lambda_{m,k}) U_k, \tag{5}$$

where U_k is an orthogonal matrix and $\lambda_{1,k} \geq \lambda_{2,k} \geq \cdots \geq \lambda_{m,k}$. Then, from (4) we have

$$|\lambda_{i,k}| - \lambda_{i,k} \leq \tau_k^{1/2}. \tag{6}$$

From (6), we deduce that

$$\lambda_{i,k} \geq -\tau_k^{1/2}/2, \quad i = 1, \cdots, m.$$

It follows that

$$g(x_k) \succeq -\tau_k^{1/2}/2I.$$

Thus, we have from the definition of $v(u)$ that

$$v(-\tau_k^{1/2}/2) \leq f(x_k).$$

This, combined with (3), yields

$$v(-\tau_k^{1/2}/2) \leq v(0) - \delta/2.$$

Hence,

$$v(0) \leq \liminf_{l \to +\infty} \leq v(0) - \delta/2,$$

which is impossible.

Necessity. Suppose to the contrary that there exist $u_k \to 0^+$ and $K > 0$ such that

$$v(u_k) \leq v(0) - \delta, \quad k \geq K.$$

As a result, there exists x_k such that

$$g(x_k) + u_k I \succeq 0 \tag{7}$$

and

$$f(x_k) \leq v(0) - \delta/2, \quad k \geq K. \tag{8}$$

Let $r_k = 1/u_k$. It follows that

$$v_1(r_k) \leq f(x_k) + 1/u_k \| |g(x_k)| - g(x_k)\|^2, \forall k.$$

This, together with (8), gives us

$$v_1(1/u_k) \leq v(0) - \delta/2 + 1/u_k \| |g(x_k)| - g(x_k)\|^2. \tag{9}$$

Assume $g(x_k)$ as in (5). Then from (7), we have

$$\lambda_{i,k} + u_k \geq 0, \quad i = 1, \cdots, m.$$

As a result,

$$0 \leq |\lambda_{i,k}| - \lambda_{i,k} \leq 2u_k.$$

This, together with (9), implies

$$\begin{aligned} v_1(1/u_k) &\leq v(0) - \delta/2 + 1/u_k \cdot 2m \cdot u_k^2 \\ &= v(0) + 2m \cdot u_k - \delta/2. \end{aligned}$$

Passing to the limit, we get

$$v(0) = \lim_{k \to +\infty} v_1(1/u_k) \leq v(0) - \delta/2,$$

which is impossible. The proof is complete. □

Some sufficient conditions that guarantee the lower semicontinuity of the perturbation function $v(u)$ at the origin are presented in the following proposition, whose proof is similar to that of Proposition 3.2 in [32].

Let $X(u) = \{x \in R^n : g(x) + uI \succeq 0\}, \quad u \geq 0.$

Proposition 3.1. Consider (SDP) and its perturbed problem (SDP_u) $(u \geq 0)$. If one of the following conditions holds, then the perturbation function $v(u)$ is lower semicontinuous at the origin.

(i) The set-valued map $X(u)$ is upper semicontinuous at $u = 0$ and $X(0) = X_0$ is compact.

(ii) The set-valued map $X(u)$ is upper semicontinuous at $u = 0$ and there exists a neighbourhood U of $X(0) = X_0$ such that f is uniformly continuous on U.

(iii) f is level-bounded on R^n, i.e., $\lim\limits_{\|x\| \to +\infty} f(x) = +\infty$.

(iv) There exists $\alpha > 0$ such that f is level-bounded on the set

$$\Lambda_\alpha = \{x \in R^n : g(x) + \alpha I \succeq 0\},$$

namely, for any sequence $\{x_k\} \subset \Lambda_\alpha$ with $\|x_k\| \to +\infty$, we have $\lim_{k \to +\infty} f(x_k) = +\infty$.

Remark 3.1. Some sufficient conditions, which are easy to verify, that guarantee the upper semicontinuity of the set-valued map $X(u)$ at the origin can be found in [33].

Denote by S and S_r^1, the sets of optimal solutions of (SDP), (PP_r), respectively.

The next theorem gives some sufficient conditions for the existence of optimal solutions to (PP_r) and their convergence.

Theorem 3.2. Consider problems (SDP) and (PP_r). Assume that (2) holds. Suppose that one of the conditions (i), (iii) and (iv) of Proposition 3.1 holds. Then

(a) S is nonempty and compact;

(b) there exists $\bar{r}' > 0$ such that S_r^1 is nonempty and compact whenever $r \geq \bar{r}'$;

(c) suppose that $x_r \in S_r^1$. Then $\{x_r\}$ is bounded and every limit point of $\{x_r\}$ belongs to S.

Proof. We only prove the case when (iv) of Proposition 3.1 holds since the other two cases are easier to prove.

(a) Since $X_0 \neq \emptyset$, fix an $x_0 \in X_0$. Then the set $\{x \in X_0 : f(x) \leq f(x_0)\} \subset \Lambda_\alpha \bigcap \{x \in R^n : f(x) \leq f(x_0)\}$ is compact. Therefore, $S \neq \emptyset$. As $S \subset \{x \in X_0 : f(x) \leq f(x_0)\}$, it follows that S is bounded. It is obvious that S is closed. Hence, S is nonempty and compact.

(b) Let $x_0 \in X_0$. We show that there exists $\bar{r}' > 0$ such that, for any $r \geq \bar{r}'$,

$$\{x \in R^n : f(x) + r\||g(x)| - g(x)\|^2 \leq f(x_0) + r\||g(x_0)| - g(x_0)\|^2 = f(x_0)\} \subset \Lambda_\alpha. \quad (10)$$

Otherwise, there exists $0 < r_k \to +\infty$ and $x_k \in R^n$ such that

$$f(x_k) + r_k\||g(x_k)| - g(x_k)\|^2 \leq f(x_0) \quad (11)$$

and

$$g(x_k) + \alpha I \not\succeq 0. \quad (12)$$

From (11) and (2), we have

$$\||g(x_k)| - g(x_k)\|^2 \leq \frac{f(x_0) - m_0}{r_k - \bar{r}} = \tau_k.$$

By the same argument as in the proof of the sufficiency part of Theorem 2.1, we have

$$g(x_k) \succeq -\tau_k^{1/2}/2I.$$

Consequently,

$$g(x_k) + \tau_k^{1/2}/2I \succeq 0$$

when k is sufficiently large. This contradicts (12). Hence, there exists $\bar{r}' > 0$ such that (10) holds. As a result, S_r^1 is nonempty and compact whenever $r \geq \bar{r}'$.

(c) Let $x_r \in S_r^1$, $r \geq \bar{r}'$. Then $\{x_r\} \subset \Lambda_\alpha$. Hence, $\{x_r\}$ is bounded. Suppose that $\bar{x}$ is a limit point of $\{x_r\}$. Then there exist $0 < r_k \to +\infty$ and $x_{r_k} \in S_{r_k}$ such that $\lim_{k\to+\infty} x_{r_k} = \bar{x}$. Let $x_0 \in X_0$. Then, from $x_{r_k} \in S_{r_k}$, we have

$$f(x_{r_k}) + r_k\||g(x_{r_k})| - g(x_{r_k})\|^2 \leq f(x_0). \quad (13)$$

It follows that

$$\||g(x_{r_k})| - g(x_{r_k})\|^2 \leq \frac{f(x_0) - f(x_{r_k})}{r_k}.$$

Passing to the upper limit as $k \to +\infty$, we obtain

$$\||g(\bar{x})| - g(\bar{x})\| \leq 0.$$

Hence, $g(\bar{x}) \succeq 0$, i.e., $\bar{x} \in X_0$. Furthermore, from (13), we have

$$f(x_{r_k}) \leq f(x_0).$$

Passing to the limit as $k \to +\infty$, we have $f(\bar{x}) \le f(x_0)$. By arbitrariness of $x_0 \in X_0$, we see that $\bar{x} \in S$. The proof is complete. □

Recall that $v_1(r)$ is the optimal value of problem (PP_r) . We have the following convergence result for approximate optimal solutions of (PP_r). The proof is elementary and thus omitted.

Theorem 3.3. Suppose that $0 < \delta_k \to 0$. Let $0 < r_k \to +\infty$ and each x_k satisfy

$$f(x_k) + r_k \| |g(x_k)| - g(x_k) \|^2 \le v_1(r_k) + \delta_k.$$

Then each limit point of $\{x_k\}$ is a solution to (SDP).

3.2 Penalty Problems $(PP_r^{\epsilon_r})$

In this subsection, we deal with the convergence of optimal values and optimal solutions of $(PP_r^{\epsilon_r})$.

Theorem 3.4. Assume that (2) holds. Then, the following two statements are true:

(i) If $\liminf_{u \to 0^+} v(u) = v(0)$, then for any sequence $0 < \epsilon_r$ with $r\epsilon_r^2 \to 0$ as $r \to +\infty$, there holds $\lim_{r \to +\infty} v_2(r, \epsilon_r) = v(0)$.

(ii) The converse of (i) is also true.

Proof. (i) Let $x_0 \in X_0$. Then

$$v_2(r, \epsilon_r) \le f(x_0) + r\text{trace}\left[(\sqrt{g^2(x_0) + \epsilon_r^2 I} - g(x_0))^2 \right].$$

It follows that

$$\begin{aligned}\limsup_{r \to +\infty} v_2(r, \epsilon_r) &\le f(x_0) + \limsup_{r \to +\infty} r\epsilon_r^2 \text{trace}\left(\frac{\epsilon_r^2 I}{(\sqrt{g^2(x_0) + \epsilon_r^2 I} + g(x_0))^2} \right) \\ &= f(x_0).\end{aligned}$$

Hence,

$$\limsup_{r \to +\infty} v_2(r, \epsilon_r) \le v(0). \tag{14}$$

Suppose to the contrary that for some $\delta > 0$,

$$\limsup_{r \to +\infty} v_2(r, \epsilon_r) \le v(0) - \delta.$$

Then there exist $r_k \to +\infty$ and $\epsilon_{r_k} > 0$ satisfying $r_k \epsilon_{r_k}^2 \to 0$ such that

$$v_1(r_k, \epsilon_{r_k}) \leq v(0) - \delta/2, \forall k.$$

It follows that there exists x_k such that

$$\begin{aligned} & m_0 + r_k \left\| (\sqrt{g^2(x_k) + \epsilon_{r_k}^2 I} - g(x_k)) \right\|^2 - \bar{r} \| |g(x_k)| - g(x_k) \|^2 \\ & \leq f(x_k) + r_k \left\| (\sqrt{g^2(x_k) + \epsilon_{r_k}^2 I} - g(x_k)) \right\|^2 \\ \leq & \ v(0) - \delta/4, \quad \forall k. \end{aligned} \tag{15}$$

As a result,

$$r_k \| \sqrt{g^2(x_k) + \epsilon_{r_k}^2 I} - g(x_k) \|^2 - \bar{r} \| |g(x_k)| - g(x_k) \|^2 \leq v(0) - m_0 - \delta/4. \tag{16}$$

Assume $g(x_k)$ as in (5). Then, from (16) we have

$$r_k \sum_{i=1}^{m} \left(\sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} - \lambda_{i,k} \right)^2 - \bar{r} \sum_{i=1}^{m} (|\lambda_{i,k}| - \lambda_{i,k})^2 \leq v(0) - m_0 - \delta/4.$$

It follows that

$$\begin{aligned} & \sum_{i=1}^{m} \left[\lambda_{i,k}^2 - \lambda_{i,k} \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right] \\ \leq & \ \frac{v(0) - m_0 - \delta/4 - m r_k \epsilon_{r_k}^2 - 2 \sum_{i=1}^{m} \bar{r} \lambda_{i,k} | \left(\lambda_{i,k} | - \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right)}{2(r_k - \bar{r})} \end{aligned}$$

Note that

$$r_k \epsilon_{r_k}^2 \to 0$$

and

$$\lambda_{i,k} \left(|\lambda_{i,k}| - \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right) \to 0, \quad i = 1, \cdots, m. \tag{17}$$

Consequently,

$$\begin{aligned} & \sum_{i=1}^{m} \left[\lambda_{i,k}^2 - \lambda_{i,k} \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right] \\ \leq & \ s_k =: \frac{v(0) - m_0 - \delta/4 - m r_k \epsilon_{r_k}^2 - 2 \sum_{i=1}^{m} \bar{r} \lambda_{i,k} | \left(\lambda_{i,k} | - \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right)}{2(r_k - \bar{r})} \to 0. \end{aligned}$$

That is,

$$\begin{aligned} & \sum_{i=1}^{m} \left[\lambda_{i,k}^2 - \lambda_{i,k} \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right] \\ & = \sum_{i=1}^{m} \lambda_{i,k} (\lambda_{i,k} - |\lambda_{i,k}|) + \sum_{i=1}^{m} \lambda_{i,k} \left[|\lambda_{i,k}| - \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right] \\ & = \sum_{\lambda_{i,k} < 0} 2 \lambda_{i,k}^2 + \sum_{i=1}^{m} \lambda_{i,k} \left[|\lambda_{i,k}| - \sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} \right] \leq s_k. \end{aligned}$$

This, combined with (17), shows that there exists a positive sequence $s'_k \to 0$ such that

$$\sum_{\lambda_{i,k}<0} 2\lambda_{i,k}^2 \leq s'_k.$$

Hence,

$$\lambda_{i,k} \geq -\sqrt{s'_k}, \quad \text{if } \lambda_{i,k} < 0.$$

So

$$\lambda_{i,k} \geq -\sqrt{s'_k}, \quad i = 1, \cdots, m.$$

Let $0 < \tau_k = \sqrt{s'_k} \to 0$. Then,

As a result,

$$g(x_k) + \tau_k I \succeq 0.$$

By assumption, we have

$$\liminf_{k\to+\infty} f(x_k) \geq \liminf v(\tau_k) = v(0). \tag{18}$$

On the other hand, from (15) we have

$$f(x_k) \leq v(0) - \delta/4.$$

It follows that

$$\limsup_{k\to+\infty} f(x_k) \leq v(0) - \delta/4,$$

contradicting (18). Hence,

$$\liminf_{r\to+\infty} v(r, \epsilon_r) \geq v(0).$$

This combined with (14) yields

$$\lim_{r\to+\infty} v(r, \epsilon_r) = v(0).$$

(ii) Suppose to the contrary that

$$\liminf_{u\to 0^+} v(u) \leq v(0) - \delta,$$

for some $\delta > 0$. Then there exists $u_k \to 0^+$ and $K > 0$ such that

$$v(u_k) \leq v(0) - \delta/2, \quad k \geq K.$$

As a result, there exists x_k such that

$$g(x_k) + u_k I \succeq 0 \tag{19}$$

and

$$f(x_k) \le v(0) - \delta/4, \quad k \ge K. \tag{20}$$

Let $r_k = 1/u_k$, $\epsilon_{r_k} = u_k$. Then $0 < r_k \to +\infty$, $\epsilon_{r_k} > 0$ and $r_k\epsilon_{r_k}^2 = u_k \to 0$. It follows that

$$v_1(r_k, \epsilon_{r_k}) \le f(x_k) + 1/u_k \text{trace}\left((\sqrt{g^2(x_k) + u_k^2 I} - g(x_k))^2\right), \forall k. \tag{21}$$

Assume $g(x_k)$ as in (5). Then

$$\text{trace}\left((\sqrt{g^2(x_k) + u_k^2 I} - g(x_k))^2\right) = \sum_{i=1}^{m}\left(\sqrt{\lambda_{i,k}^2 + u_k^2} - \lambda_{i,k}\right)^2. \tag{22}$$

From (19), we have

$$\lambda_{i,k} \ge -u_k, \quad i = 1, \cdots, m. \tag{23}$$

Consider the function

$$F(y) = \sqrt{y^2 + u_k^2} - y, y \in R.$$

We have

$$F'(y) = \frac{y}{\sqrt{y^2 + u_k^2}} - 1 = \frac{y - \sqrt{y^2 + u_k^2}}{\sqrt{y^2 + u_k^2}} < 0.$$

Hence, $F(y)$ is decreasing. In addition, $F(y) > 0, \forall y \in R$. These properties of $F(y)$ combined with (21)-(23) yield

$$v_1(r_k, \epsilon_{r_k}) \le f(x_k) + mu_k(\sqrt{2} + 1)^2.$$

This, together with (20), gives us

$$v_1(r_k, \epsilon_{r_k}) \le v(0) - \delta/4 + mu_k(\sqrt{2} + 1)^2.$$

As a result,

$$\limsup_{k \to +\infty} v_1(r_k, \epsilon_{r_k}) \le v(0) - \delta/4,$$

contradicting the assumption. The proof is complete. □

Denote by S_r^2 the set of optimal solutions of $(PP_r^{\epsilon_r})$.

The next theorem gives sufficient conditions for the existence of optimal solutions to $(PP_r^{\epsilon_r})$ and their convergence.

Theorem 3.5. Consider problems (SDP) and $(PP_r^{\epsilon_r})$. Assume that (2) holds. Suppose that one of the conditions (i), (iii) and (iv) of Proposition 3.1 holds. Then

(a) S is nonempty and compact;

(b) there exists $\bar{r}'' > 0$ such that S_r^2 is nonempty and compact whenever $r \ge \bar{r}''$;

(c) suppose that $x_r \in S_r^2$, $r \geq \bar{r}''$. Then $\{x_r\}$ is bounded and every limit point of $\{x_r\}$ belongs to S.

Proof. We only prove the case when (iv) of Proposition 3.1 holds since the other two cases are easier to prove.

(a) The same as the proof of statement (a) of Theorem 3.2.

(b) Let $x_0 \in X_0$. We show that there exists $\bar{r}'' > 0$ such that for $r \geq \bar{r}''$,

$$\{x \in R^n : f(x) + r\|\sqrt{g^2(x) + \epsilon_r^2 I} - g(x)\|^2 \leq f(x_0) + r\|\sqrt{g^2(x_0) + \epsilon_r^2 I} - g(x_0)\|^2\} \subset \Lambda_\alpha. \quad (24)$$

Otherwise, there exists $0 < r_k \to +\infty$ and $0 < \epsilon_{r_k}$ with $r_k \epsilon_{r_k}^2 \to 0$, and $x_k \in R^n$ such that

$$f(x_k) + r_k\|\sqrt{g^2(x_k) + \epsilon_{r_k}^2 I} - g(x_k)\|^2 \leq f(x_0) + r_k\|\sqrt{g^2(x_0) + \epsilon_{r_k}^2 I} - g(x_0)\|^2 \quad (25)$$

and

$$g(x_k) + \alpha I \not\succeq 0. \quad (26)$$

From (25) and (2), we have

$$\begin{aligned} &r_k\|\sqrt{g^2(x_k) + \epsilon_{r_k}^2 I} - g(x_k)\|^2 - \bar{r}\||g(x_k)| - g(x_k)\|^2 \\ &\leq f(x_0) - m_0 + \|\sqrt{g^2(x_0) + \epsilon_{r_k}^2 I} - g(x_0)\|^2. \end{aligned}$$

Arguing as in the proof of the sufficiency part of Theorem 3.4, there exist a subsequence $\{r_{k_l}\}$ of $\{r_k\}$ and a sequence $0 < \tau_l \to 0$ such that

$$g(x_{k_l}) + \tau_l I \succeq 0.$$

Consequently,

$$g(x_{k_l}) + \alpha I \succeq 0$$

when l is sufficiently large. This contradicts (26). Hence, there exists $\bar{r} > 0$ such that (24) holds. As a result, S_r^2 is nonempty and compact whenever $r \geq \bar{r}''$.

(c) Let $x_r \in S_r^2$, $r \geq \bar{r}''$. Then $\{x_r\} \subset \Lambda_\alpha$. Hence, $\{x_r\}$ is bounded. Suppose that $\bar{x}$ is a limit point of $\{x_r\}$. Then there exist $0 < r_k \to +\infty$ and $x_{r_k} \in S_{r_k}^2$ such that $\lim_{k\to+\infty} x_{r_k} = \bar{x}$. Let $x_0 \in X_0$. Then from $x_{r_k} \in S_{r_k}^2$, we have

$$f(x_{r_k}) + r_k\|\sqrt{g^2(x_{r_k}) + \epsilon_{r_k}^2 I} - g(x_{r_k})\|^2 \leq f(x_0) + r_k\|\sqrt{g^2(x_0) + \epsilon_{r_k}^2 I} - g(x_0)\|^2. \quad (27)$$

It follows that

$$\|\sqrt{g^2(x_{r_k}) + \epsilon_{r_k}^2 I} - g(x_{r_k})\|^2 \leq \frac{f(x_0) - f(x_{r_k})}{r_k} + \|\sqrt{g^2(x_0) + \epsilon_{r_k}^2 I} - g(x_0)\|^2.$$

Passing to the upper limit as $k \to +\infty$, we obtain

$$\||g(\bar{x})| - g(\bar{x})\| \le 0.$$

Hence, $g(\bar{x}) \succeq 0$, i.e., $\bar{x} \in X_0$. Furthermore, from (27), we have

$$\begin{aligned} f(x_{r_k}) &\le f(x_0) + r_k \left\| \sqrt{g^2(x_0) + \epsilon_{r_k}^2 I} - g(x_0) \right\|^2 \\ &= f(x_0) + r_k \left\| \frac{\epsilon_{r_k}^2 I}{\sqrt{g^2(x_0) + \epsilon_{r_k}^2 I} + g(x_0)} \right\|^2 \\ &= f(x_0) + r_k \epsilon_{r_k}^2 \left\| \frac{\epsilon_{r_k} I}{\sqrt{g^2(x_0) + \epsilon_{r_k}^2 I} + g(x_0)} \right\|^2. \end{aligned} \tag{28}$$

Note that $r_k \epsilon_{r_k}^2 \to 0$ and $\left\{ \left\| \frac{\epsilon_{r_k} I}{\sqrt{g^2(x_0)+\epsilon_{r_k}^2 I} + g(x_0)} \right\|^2 \right\}$ is bounded. Taking the limit in (28) as $k \to +\infty$, we have $f(\bar{x}) \le f(x_0)$. By the arbitrariness of $x_0 \in X_0$, we see that $\bar{x} \in S$. The proof is complete. □

Recall that $v_2(r, \epsilon_r)$ is the optimal value of problem $(PP_r^{\epsilon_r})$. We have the following convergence result for the approximate optimal solutions of $(PP_r^{\epsilon_r})$, whose proof is similar to that of Theorem 3.3 and thus omitted.

Theorem 3.6. Suppose that $0 < \delta_k \to 0$. Let $0 < r_k \to +\infty$ and $0 < \epsilon_{r_k}$ satisfy $r_k \epsilon_{r_k}^2 \to 0$. Let each x_k satisfy

$$f(x_k) + r_k \|\sqrt{g^2(x_k) + \epsilon_{r_k}^2 I} - g(x_k)\|^2 \le v_1(r_k, \epsilon_{r_k}) + \delta_k.$$

Then each limit point of $\{x_k\}$ is a solution to (SDP).

4 Convergence of Stationary Points of the Penalty Problems

In this section, we present necessary optimality conditions for a local minimum of (PP_r) ($(PP_r^{\epsilon_r})$). We show that any limit point of a sequence of stationary points of (PP_r) ($(PP_r^{\epsilon_r})$) satisfies the KKT optimality condition of (SDP).

Definition 4.1 [23]. Let $x_0 \in R^n$ be feasible to (SDP). We say that the Mangasarian-Fromovitz constraint qualification holds at x_0 if there exists $d \in R^n$ such that $g(x_0) +$

$Dg(x_0)(d) \succ 0$.

Definition 4.2. Let $\bar{x}$ be feasible to (SDP). We say that $\bar{x}$ satisfies the KKT optimality condition of (SDP) if there exists $\Omega \in S_m$ with $\Omega \succeq 0$ such that

$$\frac{\partial f(\bar{x})}{\partial x_i} - \text{trace}\left(\Omega \frac{\partial g(\bar{x})}{\partial x_i}\right) = 0 \tag{29}$$

and

$$\Omega g(\bar{x}) = 0. \tag{30}$$

It was established in [23] that if $\bar{x}$ is a local solution of (SDP) and the Mangasarian-Fromovitz constraint qualification holds at $\bar{x}$. Then $\bar{x}$ is a KKT point of (SDP).

First we give necessary optimality conditions for (PP_r).

Theorem 4.1. Suppose that $\bar{x}_r$ is a local minimum of (PP_r). Then

$$\frac{\partial f(\bar{x}_r)}{x_i} + r\text{trace}\left[(g(\bar{x}_r) - |g(\bar{x}_r)|)\frac{\partial g(\bar{x}_r)}{\partial x_i}\right] = 0, \quad i = 1, \cdots, n. \tag{31}$$

Proof. The conclusion follows directly from Proposition 2.1 and the standard necessary optimality conditions for a local minimum of an unconstrained optimization problem. □

Let

$$h(x) = \|\sqrt{g^2(x) + \epsilon_r^2 I} - g(x)\|^2.$$

It is straightforward to prove the next lemma.

Lemma 4.1.

$$\frac{\partial h(x)}{\partial x_i} = 2\text{trace}\left[\left(2g(x) - \frac{2g^2(x) + \epsilon_r^2 I}{\sqrt{g^2(x) + \epsilon_r^2 I}}\right)\frac{\partial g(x)}{\partial x_i}\right], \quad i = 1, \cdots, n. \tag{32}$$

Now we derive optimality conditions for a local minimum of $(PP_r^{\epsilon_r})$.

Theorem 4.2. Let $\bar{x}_r$ be a local solution to $(PP_r^{\epsilon_r})$. Then

$$\frac{\partial f(\bar{x}_r)}{x_i} + 2r\text{trace}\left[\left(2g(\bar{x}_r) - \frac{2g^2(\bar{x}_r) + \epsilon_r^2 I}{\sqrt{g^2(\bar{x}_r) + \epsilon_r^2 I}}\right)\frac{\partial g(\bar{x}_r)}{\partial x_i}\right] = 0, \quad i = 1, \cdots, n. \tag{33}$$

Proof. Since $\bar{x}_r$ is a local solution to $(PP_r^{\epsilon_r})$, by the standard necessary optimality condition, we have

$$\frac{\partial f(\bar{x}_r)}{x_i} + r\frac{\partial h(\bar{x}_r)}{\partial x_i} = 0, \qquad i = 1, \cdots, n. \tag{34}$$

Substituting (32) into (34), we obtain (33). □

4.1 Penalty Problems (PP_r)

In this subsection, we show the convergence of stationary points of (PP_r).

The next lemma is useful for convergence analysis. Since the proof is straightforward, we omit it.

Lemma 4.2 Let $0 < r_k \to +\infty$. Let $\bar{x}_k \in R^n, \forall k$. Suppose that there exists $M \in R$ such that

$$f(\bar{x}_k) + r_k \| |g(\bar{x}_k)| - g(\bar{x}_k) \|^2 \leq M. \tag{35}$$

Then any limit point of $\{\bar{x}_k\}$ is feasible to (SDP).

The convergence of stationary points of (PP_r) is presented in the following theorem.

Theorem 4.3. Let $0 < r_k \to +\infty$ as $k \to +\infty$. Consider the problems (SDP) and (PP_{r_k}). Let each $\bar{x}_k$ be generated by some method for solving (PP_{r_k}). Suppose that there exists $M \in R$ such that (35) holds. Then each limit point of $\{\bar{x}_k\}$ is feasible for (SDP). Furthermore, suppose that each $\bar{x}_k$ satisfies the optimality condition of (PP_{r_k}) given by (31) (with r replaced by r_k). Let $\bar{x}$ be a limit point of $\{\bar{x}_k\}$ and let the Mangasarian-Fromovitz constraint qualification hold at $\bar{x}$. Then $\bar{x}$ satifies the KKT optimality condition of (SDP).

Proof. By Lemma 4.2, each limit point of $\{\bar{x}_k\}$ is feasible for (SDP). Assume without loss of generality that $\bar{x}_k \to \bar{x}$ as $k \to +\infty$. Let

$$\Omega_k = -2r_k \left[g(\bar{x}_k) - |g(\bar{x}_k)| \right] \succeq 0. \tag{36}$$

Then (31) (with r replaced by r_k) becomes

$$\frac{\partial f(\bar{x}_k)}{\partial x_i} - \text{trace} \left[\Omega_k \frac{\partial g(\bar{x}_k)}{\partial x_i} \right] = 0, \quad i = 1, \cdots, n. \tag{37}$$

We assert that $\{\Omega_k\}$ is bounded. Otherwise, assume without loss of generality that $\|\Omega_k\| \to +\infty$ and

$$\lim_{k \to +\infty} \Omega_k / \|\Omega_k\| = \Omega' \succeq 0.$$

Dividing (37) by $\|\Omega_k\|$ and passing to the limit as $k \to +\infty$, we get

$$\text{trace} \left(\Omega' \frac{\partial g(\bar{x})}{\partial x_i} \right) = 0, \quad i = 1, \cdots, n. \tag{38}$$

Note that

$$\begin{aligned}\operatorname{trace}(\Omega' g(\bar{x})) &= \lim_{k\to+\infty} \operatorname{trace}\left(\frac{\Omega_k}{\|\Omega_k\|} g(\bar{x}_k)\right) \\ &= \lim_{k\to+\infty} \operatorname{trace}\left(\frac{|g(\bar{x}_k)|(g(\bar{x}_k) - |g(\bar{x}_k)|)}{\|(g(\bar{x}_k) - |g(\bar{x}_k)|\|}\right) \\ &\le 0\end{aligned}$$

because $|g(\bar{x}_k)| \succeq 0$ and $g(\bar{x}_k) - |g(\bar{x}_k)| \preceq 0$.

On the other hand, from $\Omega' \succeq 0$ and $g(\bar{x}) \succeq 0$, we deduce that

$$\operatorname{trace}(\Omega' g(\bar{x})) \ge 0.$$

Hence, we have

$$\operatorname{trace}(\Omega' g(\bar{x})) = 0. \tag{39}$$

By the Mangasarian-Fromovitz constraint qualification at $\bar{x}$, there exists $d \in R^n$ such that $g(\bar{x}) + Dg(\bar{x})(d) \succ 0$. It is obvious that $\Omega' \neq 0$. It follows that

$$\operatorname{trace}\left(\Omega'(g(\bar{x}) + Dg(\bar{x})(d)\right) > 0.$$

This, combined with (39), yields

$$\operatorname{trace}\left(\Omega' Dg(\bar{x})(d)\right) > 0,$$

contradicting (38). So we assume without loss of generality that $\Omega_k \to \Omega \succeq 0$. Taking the limit in (37) as $k \to +\infty$, we obtain (29). Moreover,

$$\begin{aligned}\operatorname{trace}(\Omega g(\bar{x})) &= \lim_{k\to+\infty} \operatorname{trace}\left(\Omega_k g(\bar{x}_k)\right) \\ &= \lim_{k\to+\infty} \operatorname{trace}\left(r_k |g(\bar{x}_k)|(g(\bar{x}_k) - |g(\bar{x}_k)|)\right) \\ &\le 0\end{aligned}$$

In the meantime, $\operatorname{trace}(\Omega g(\bar{x})) \ge 0$. Hence, $\operatorname{trace}(\Omega g(\bar{x})) = 0$, implying (30). The proof is complete. □

4.2 Penalty Problems $(PP_r^{\epsilon_r})$

In this subsection, we carry out convergence analysis of the stationary points of $(PP_r^{\epsilon_r})$.

We need the following lemma, whose proof is straightforward and thus omited.

Lemma 4.3. Let $0 < r_k \to +\infty$ and $0 < \epsilon_{r_k} \to 0$. Let $\bar{x}_k \in R^n, \forall k$. Suppose that there exists $M \in R$ such that

$$f(\bar{x}_k) + r_k \|\sqrt{g^2(\bar{x}_k) + \epsilon_{r_k}^2 I} - g(\bar{x}_k)\|^2 \le M. \tag{40}$$

Then any limit point of $\{\bar{x}_k\}$ is feasible to (SDP).

The following theorem gives convergence results for the stationary points of the penalty problems $(PP_r^{\epsilon_r})$.

Theorem 4.4. Let $0 < r_k \to +\infty$ and $0 < r_k \epsilon_{r_k}^2 \to 0$ as $k \to +\infty$. Consider the problems (SDP) and $(PP_{r_k}^{\epsilon_{r_k}})$. Let each $\bar{x}_k$ be generated by some method for solving $(PP_{r_k}^{\epsilon_{r_k}})$. Suppose that there exists $M \in R$ such that (40) holds. Then each limit point of $\{\bar{x}_k\}$ is a feasible solution to (SDP). Furthermore, suppose that each $\bar{x}_k$ satisfies the optimality condition of $(PP_{r_k}^{\epsilon_{r_k}})$ given by (33) (with r and ϵ_r replaced by r_k and ϵ_{r_k}, respectively). Let $\bar{x}$ be a limit point of $\{\bar{x}_k\}$ and let the Mangasarian-Fromovitz constraint qualification hold at $\bar{x}$. Then $\bar{x}$ satifies the KKT optimality condition of (SDP).

Proof. The assertion that each limit point of $\{\bar{x}_k\}$ is a feasible solution to (SDP) follows directly from Lemma 4.3. Assume without loss of generality that $\bar{x}_k \to \bar{x}$ as $k \to +\infty$. Let

$$\Omega_k = -2r_k \left[2g(\bar{x}_k) - \frac{2g^2(\bar{x}_k) + \epsilon_{r_k}^2 I}{\sqrt{g^2(\bar{x}_k) + \epsilon_{r_k}^2 I}} \right]. \tag{41}$$

Then (33) (with r and ϵ_r replaced by r_k and ϵ_{r_k}, respectively) becomes

$$\frac{\partial f(\bar{x}_k)}{\partial x_i} - \text{trace} \left[\Omega_k \frac{\partial g(\bar{x}_k)}{\partial x_i} \right] = 0, \quad i = 1, \cdots, n. \tag{42}$$

Note that

$$\frac{2g^2(\bar{x}_k) + \epsilon_{r_k}^2 I}{\sqrt{g^2(\bar{x}_k) + \epsilon_{r_k}^2 I}} = \sqrt{g^2(\bar{x}_k) + \epsilon_{r_k}^2 I} + \frac{g^2(\bar{x}_k)}{\sqrt{g^2(\bar{x}_k) + \epsilon_{r_k}^2 I}} \succeq 2g(\bar{x}_k).$$

Consequently,

$$\Omega_k \succeq 0. \tag{43}$$

Now we prove that $\{\Omega_k\}$ is bounded. Otherwise, assume without loss of generality that $\|\Omega_k\| \to +\infty$ and

$$\lim_{k \to +\infty} \Omega_k / \|\Omega_k\| = \Omega'. \tag{44}$$

It is clear from (43) and (44) that

$$\Omega' \succeq 0, \tag{45}$$

$$\|\Omega'\| = 1. \tag{46}$$

Dividing (42) by $\|\Omega_k\|$, we obtain

$$\frac{\frac{\partial f(\bar{x}_k)}{\partial x_i}}{\|\Omega_k\|} - \text{trace}\left[\frac{\Omega_k}{\|\Omega_k\|}\frac{\partial g(\bar{x}_k)}{\partial x_i}\right] = 0, \quad i = 1, \cdots, n. \tag{47}$$

Taking the limit in (47) as $k \to +\infty$, we get

$$\text{trace}\left(\Omega' \frac{\partial g(\bar{x})}{\partial x_i}\right) = 0, \quad i = 1, \cdots, n. \tag{48}$$

Assume that

$$g(\bar{x}_k) = U_k^T \text{diag}(\lambda_{1,k}, \cdots, \lambda_{m,k}) U_k,$$

where U_k is an $m \times m$ orthogonal matrix and $\lambda_{1,k} \geq \lambda_{2,k} \geq \cdot \geq \lambda_{m,k}$. As a result,

$$\begin{aligned}\Omega_k g(\bar{x}_k) = 2r_k U_k^T \text{diag}&\left(\frac{\left(\sqrt{\lambda_{1,k}^2 + \epsilon_{r_k}^2} - \lambda_{1,k}\right)^2}{\sqrt{\lambda_{1,k}^2 + \epsilon_{r_k}^2}}\lambda_{1,k}, \cdots,\right.\\ &\left.\frac{\left(\sqrt{\lambda_{m,k}^2 + \epsilon_{r_k}^2} - \lambda_{m,k}\right)^2}{\sqrt{\lambda_{m,k}^2 + \epsilon_{r_k}^2}}\lambda_{m,k}\right) U_k.\end{aligned} \tag{49}$$

Assume that

$$g(\bar{x}) = U^T \text{diag}\left(\lambda_1, \cdots, \lambda_s, 0 \cdots, 0\right) U, \tag{50}$$

where U is an $m \times m$ orthogonal matrix, $s = \text{rank}(g(\bar{x}))$ and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_s > 0$.

By the continuity of g, we have

$$\lim_{k \to +\infty} \lambda_{i,k} = \lambda_i > 0, \quad i = 1, \cdots, s, \tag{51}$$

$$\lim_{k \to +\infty} \lambda_{i,k} = 0, \quad i = s+1, \cdots, m. \tag{52}$$

Therefore,

$$\begin{aligned}&\lim_{k \to +\infty} 2r_k \frac{\left(\sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} - \lambda_{i,k}\right)^2}{\sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2}}\lambda_{i,k}\\ &= 2\lim_{k \to +\infty} \frac{r_k \epsilon_{r_k}^4}{\sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2}} \cdot \left(\sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} + \lambda_{i,k}\right)^2 \lambda_{i,k}\\ &= 0 \cdot \frac{1}{4\lambda_i^3} = 0, \quad i = 1, \cdots, s.\end{aligned} \tag{53}$$

Hence,

$$\begin{aligned}&\lim_{k \to +\infty} 2r_k \frac{\left(\sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2} - \lambda_{i,k}\right)^2}{\sqrt{\lambda_{i,k}^2 + \epsilon_{r_k}^2}}\lambda_{i,k}/\|\Omega_k\|\\ &= 0, \quad i = 1, \cdots, s.\end{aligned} \tag{54}$$

On the other hand,

$$\lim_{k\to+\infty}\left|2r_k\frac{\left(\sqrt{\lambda_{i,k}^2+\epsilon_{r_k}^2}-\lambda_{i,k}\right)^2}{\sqrt{\lambda_{i,k}^2+\epsilon_{r_k}^2}}\lambda_{i,k}/\|\Omega_k\|\right|$$
$$\leq \lim_{k\to+\infty}|\lambda_{i,k}|=0,\quad i=s+1,\cdots,m. \tag{55}$$

The combination of (49), (54) and (55) yields

$$\lim_{k\to+\infty}\left\|\frac{\Omega_k}{\|\Omega_k\|}g(\bar{x}_k)\right\|=\|\Omega' g(\bar{x})\|=0.$$

So

$$\Omega' g(\bar{x})=0. \tag{56}$$

Since the Mangasarian-Fromovitz constraint qualification holds at $\bar{x}$, there exists $d\in R^n$ such that

$$g(\bar{x})+Dg(\bar{x})(d)\succ 0.$$

Therefore, when $t>0$ is sufficiently small,

$$g(\bar{x})+Dg(\bar{x})(d)-t\Omega'\succ 0. \tag{57}$$

It follows from (48) that

$$\text{trace}(\Omega' Dg(\bar{x})(d))=0. \tag{58}$$

(56)-(58) and (46) together give us

$$\begin{aligned}0 &\leq \text{trace}\left(\Omega'\left(g(\bar{x})+Dg(\bar{x})(d)-t\Omega'\right)\right)\\ &=\text{trace}(\Omega' g(\bar{x}))+\text{trace}(\Omega' Dg(\bar{x})(d))-t\|\Omega'\|^2\\ &=0+0-t=-t<0,\end{aligned}$$

which is impossible. So we conclude that $\{\Omega_k\}$ is bounded. Assume without loss of generality that

$$\lim_{k\to+\infty}\Omega_k=\Omega\succeq 0. \tag{59}$$

Taking the limit in (42) as $k\to+\infty$, we obtain (29). Further from (49)-(52), we can establish (53), and by the boundedness of $\{\|\Omega_k\|\}$, we have

$$\lim_{k\to+\infty}\left|2r_k\frac{\left(\sqrt{\lambda_{i,k}^2+\epsilon_{r_k}^2}-\lambda_{i,k}\right)^2}{\sqrt{\lambda_{i,k}^2+\epsilon_{r_k}^2}}\lambda_{i,k}\right|$$
$$\leq \lim_{k\to+\infty}\|\Omega_k\||\lambda_{i,k}|$$
$$=0,\quad i=s+1,\cdots,m. \tag{60}$$

The combination of (49), (53) and (60) implies

$$\Omega g(\bar{x}) = \lim_{k \to +\infty} \Omega_k g(\bar{x}_k) = 0. \tag{61}$$

(29), (59) and (61) together show that $\bar{x}$ is a KKT point of (SDP). The proof is complete. □

5 Conclusions

A nonlinear semidefinite program was converted into a mathematical program with a matrix equality constraint. A sequential quadratic penalty method was applied to the converted problem. Necessary and sufficient conditions for the convergence of optimal values of the penalty problems were given. Some sufficient conditions were provided for the existence and convergence of optimal solutions of the penalty problems. Under certain conditions, it was shown that any limit point of a sequence of stationary points of the penalty problems is a KKT stationary point of the original semidefinite programming problem.

Acknowledgement

The authors are grateful to a referee for pointing out reference [25] and providing constructive remarks on Lemmas 2.1-2.3, which help us improve the presentation of this paper.

References

[1] Alizadeh, F., Interior-point methods in semidefinite programming with applications in combinatorial problems, SIAM J. Optimization, Vol. 5, pp. 13-51, 1995.

[2] Bel-Tal, A., Jarre, F., Kocvara, M., Nemirovski and Zowe, J., Optimal design of trusses under a nonconvex global buckling constraints, Optimization and Engineering, Vol. 1, 189-213, 2000.

[3] Bonnans, J. F., Cominetti, R. and Shapiro, A., Sensitivity analysis of optimization problems under second order regular constraints, Mathematics of Operations Research, Vol. 23, pp. 806-831, 1998.

[4] Bonnans, J. F., Cominetti, R. and Shapiro, A., Second order optimality conditions based on second order tangent sets, SIAM J. Optimization, Vol. 9, pp. 466-492, 1999.

[5] Burer, S., Monteiro, R. D. C. and Zhang, Y., Solving a class of semidefinite programs via nonlinear programming, Mathematical Programming, Ser. A., Vol. 93, 97-122, 2002.

[6] Burer, S., Monteiro, R. D. C. and Zhang, Y., Interior-point algorithms for semidefinite programming based on a nonlinear formulation, Compuational Optimization and Applications, Vol. 22, pp. 49-79, 2002.

[7] Doyle, J. C., Packard, A and Zhou, K., Review of LFT's, LMI's and μ, in Proc. IEEE Conference on Decision and Control, Vol. 2, pp. 1227-1232, Brighton, Dec. 1991.

[8] Forsgren, A., Optimality conditions for nonconvex semidefinite programming, Mathematical Programming, Ser. A., Vol. 88, pp. 105-128, 2000.

[9] Fares, B., Apkarian, P. and Noll, D., Augmented Lagrangian method for a class of LMI-constrained problems in robust control theory, International Journal of Control, Vol. 74, pp. 348-360, 2001.

[10] Fiacco, A. V. and McCormick, G. P., Nonlinear Programming, Sequential Minimization Techniques, SIAM, Philadelphia, 1990.

[11] Fares, B., Noll, D. and Apkarian, P., Robust control via sequential semidefinite programming, SIAM J. Control and Optimi., Vol. 40, pp. 1791-1820, 2002.

[12] Ghaoui, L. E. and Niculescu, S. I., Advances in Linear Matrix Inequality Methods in Control, Advances in Design Control, SIAM, Pliladelphia, 2000.

[13] Horn, R. A. and Johnson, C. R., Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[14] Jarre, F., Eigenvalue problems and nonconvex minimization, in Handbook of Semidefinite Programming, Theory, Algorithms and Applications, H. Wolkowicz, R. Saigal and Vandenberghe (eds), Kluwer Academic Publishers, 2000.

[15] Jarre, F., An interior point method for semidefinite programs, Optimization and Engineering, Vol. 1, pp. 347-372, 2000.

[16] Jarre, F., Convex analysis on symmetric matrices, in Handbook of Semidefinite Programming, Theory, Algorithms and Applications, H. Wolkowicz, R. Saigal and Vandenberghe (eds), Kluwer Academic Publishers, 2000.

[17] Kanzow, C. and Nagel, C., Semidefinite programs: new search directions, smoothing-type methods, and numerical results, SIAM J. Optimization, Vol. 13, pp. 1-23, 2002.

[18] Mosheyev, L. and Zibulevsky, M., Penalty/barrier multiplier algorithm for semidefinite programming, Optimization Methods and Software, Vol. 13, pp. 235-261, 2000.

[19] Nesterov, Y., Wolkowicz, H. and Ye, Y., Semidefinite programming relaxtions of non-convex quadratic optimization, in Handbook of Semidefinite Programming, Wolkowicz, H., Saigal, R. and Vandenberghe (eds), Kluwer, Boston, 2000.

[20] Overton, M. L. and Womersley, R. S., Second deriatives for optimizing eigenvalues of symmetric matrices, SIAM J. Matrix Analysis and Applications, Vol. 16, pp. 697-718, 1995.

[21] Ringertz, U. T., Eigenvalues in optimal structural design. In: Biegler, L. T., Coleman, T. F. Conn, A. R. and Santosa, F. N. (eds), Large Scale Optimization and Applications, Part I: Optimization in Inverse Problems and Design, Vol. 92 of the IMA Volumes in Mathematics and its Applications, pp. 135-149, Springer, New York, 1997.

[22] Ringertz, U. T., Large scale structural design optimization, In: Biegler, L. T., Coleman, T. F. Conn, A. R. and Santosa, F. N. (eds), Large Scale Optimization and Applications, Part II: Optimization in Inverse Problems and Design, Vol. 92 of the IMA Volumes in Mathematics and its Applications, pp. 135-149, Springer, New York, 1997.

[23] Shapiro, A., First and second order analysis of nonlinear semidefinite programs, Mathematical Programming, Ser. B., Vol. 77, pp. 301-320, 1997.

[24] Shapiro, A., On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints, SIAM J. Optimization, Vol. 7, pp. 508-518, 1997.

[25] Shapiro, A., On differentiability of symmetric matrix valued functions, E-print available at: http://www.optimization-online.org, 2002.

[26] Shapiro, A. and Fan, M. K. H., On eigenvalue optimization, SIAM J. Optimization, Vol. 5, pp. 552-569, 1995.

[27] Sun, J., Sun, D. F. and Qi, L. Q., Quadratic convergence of a squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems, preprint, School of Business and Singapore-MIT Alliance, National University of Singapore, Singapore, 2002.

[28] Todd, M., Semidefinite Optimization, Acta Numerica, Vol. 10, pp. 515-560, 2001.

[29] Tseng, P., Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation, preprint, Department of Mathematics, University of Washington, Seattle, USA, 2002.

[30] Vandenberghe, L. and Boyd, S., Semidefinite programming, SIAM Review, Vol. 38, pp. 49-95, 1996.

[31] Wolkowicz, H., Saigal, R. and Vandenberghe, L. (eds), Handbook of semidefinite programming, theory, algorithms and applications, International Series in Operations Research and Management Science, Vol. 27, Kluwer Academic Publishers, Boston, MA, 2000.

[32] Yang, X. Q. and Huang, X. X., Partially strictly monotone and nonlinear penalty functions for constrained mathematical programs, Computational Optimization and Applications (accepted).

[33] Yang, X. Q., Wang, C. Y. and Yang, X. M., Nonlinear Lagrange duality theorems and penalty methods in continuous optimization, submitted for publication.

[34] Ye, Y., Interior Point Algorithms: Theory and Analysis, John Wiley & Son, New York, 1997.

[35] Ye, Y., Approximating quadratic programming with bound and quadratic constraints, Mathematical Programming, Vol. 84, pp. 219-226, 1999.

[36] Zhang, S. Z., Quadratic maximization and semidefinite relaxation, Mathematical Programming, Vol. 87, pp. 453-465, 2000.