Journal of Inequalities in Pure and Applied Mathematics

http://jipam.vu.edu.au/

Volume 5, Issue 2, Article 37, 2004

GLOBAL CONVERGENCE OF A MODIFIED SQP METHOD FOR MATHEMATICAL PROGRAMS WITH INEQUALITIES AND EQUALITIES CONSTRAINTS

ZHONG WAN

COLLEGE OF MATHEMATICS AND ECONOMETRICS
HUNAN UNIVERSITY, 410082
THE PEOPLE'S REPUBLIC OF CHINA.
wanmath@163.com

Received 23 December, 2003; accepted 29 April, 2004
Communicated by A. Rubinov

ABSTRACT. When we solve an ordinary nonlinear programming problem by the most and popular sequential quadratic programming (SQP) method, one of the difficulties that we must overcome is to ensure the consistence of its QP subproblems. In this paper, we develop a new SQP method which can assure that the QP subproblem at every iteration is consistent. One of the main techniques used in our method involves solving a least square problem in addition to solving a modified QP subproblem at each iteration, and we need not add bound constraints to the search direction. we also establish the global convergence of the proposed algorithm.

Key words and phrases: Nonlinear Programs, Inequalities and Equalities System, Sequential quadratic programming.

2000 *Mathematics Subject Classification.* 90C30, 90D65.

1. INTRODUCTION

We consider the following smooth nonlinear programs:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) \geq 0, h(x) = 0. \end{aligned} \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^l$, $h : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable. Among all robust methods for (1.1), the sequential quadratic programming method (SQP) is one of the most important and the most popular. The basic idea of the classical SQP is as follows: at the present iterative point x, approximate (1.1) by quadratic programs (QP) of the form:

$$\begin{aligned} \min \quad & \nabla f(x)^T d + \tfrac{1}{2} d^T B d \\ \text{s.t.} \quad & g^{'}(x)d + g(x) \geq 0, \\ & h^{'}(x)d + h(x) = 0, \end{aligned} \tag{1.2}$$

ISSN (electronic): 1443-5756

The work of this author was supported by the Postdoctor Research Funds of Central South University, P.R. China.

178-03

where $B \in \mathbb{R}^{n\times n}$ is symmetric positive definite, and $g^{'}(x) \in \mathbb{R}^{l\times n}$, $h^{'}(x) \in \mathbb{R}^{m\times n}$ are defined as follows:

$$g^{'}(x) \equiv \left(\frac{\partial g_i}{\partial x_j}\right), \quad h^{'}(x) \equiv \left(\frac{\partial h_i}{\partial x_j}\right).$$

The iteration then has the form

$$\bar{x} = x + td,$$

where d solves (1.2) and t is a step length chosen to reduce the value of some merit function for (1.1). In this paper, the merit function is taken as

$$\theta_{\rho^g,\rho^h}(x) = f(x) + \rho^g \sum_{i=1}^{l} \max\{-g_i(x), 0\} + \rho^h \|h(x)\|_2^2.$$

On one hand, one of the major priorities of SQP lies in that it does not require that the approximate solution obtained at each iteration is feasible for (1.1). On the other hand, this makes it possible that the subproblem (1.2) is not consistent. In [1], J.V. Burke and S.-P. Han describe a robust SQP wherein the QP (1.2) is altered in a way which guarantees that the associated region is nonempty for each $x \in \mathbb{R}^n$ and for which a global convergence theory is established. Recently, H. Jiang and D. Ralph developed a new modified SQP method in [3] wherein a similar global convergence result is obtained under the condition that the following modified QP

$$(1.3) \qquad \begin{aligned} \min \quad & \nabla f(x)^T d + \tfrac{1}{2} d^T B d + \rho \textstyle\sum_{i=1}^{l} s_i \\ \text{s.t.} \quad & g^{'}(x)d + g(x) \geq -s, \\ & h^{'}(x)d + h(x) = 0, \\ & s \geq 0 \end{aligned}$$

is feasible, where ρ is a penalty parameter, and s is an artificial variable. The proposed SQP method in this paper is close to [3], but removes the above condition. Our approach to guarantee the non-emptiness of constraints region of the QP subproblem comes from the ideas in [1].

2. Algorithm and its Validity

In this section, we first describe the algorithm, then we verify the validity of the proposed algorithm.

Step 0. (Initialization) Let $\rho_{-1} > 0$, $\delta_1 > 0$, $\delta_2 > 0$, $\delta_3 > 0$, $\sigma \in (0,1)$, $\tau \in (0,1)$. Choose $x^0 \in \mathbb{R}^n$ and a symmetric positive definite matrix B_0. Set $k := 0$.

Step 1. (Search direction) With $x = x^k$, solve the following linear least square problem:

$$(2.1) \qquad \min_{d\in\mathbb{R}^n} \frac{1}{2} \|h^{'}(x)d + h(x)\|_2^2.$$

Let $\tilde{d}$ be a solution of (2.1), compute $r(x) = h^{'}(x)\tilde{d} + h(x)$, and solve the following modified QP problem with $x = x^k$, $B = B_k$, $\rho = \rho_{k-1}$:

$$(2.2) \qquad \begin{aligned} \min \quad & \nabla f(x)^T d + \tfrac{1}{2} d^T B d + \rho \textstyle\sum_{i=1}^{l} s_i \\ \text{s.t.} \quad & g^{'}(x)d + g(x) \geq -s, \\ & h^{'}(x)d + h(x) = r(x), \\ & s \geq 0. \end{aligned}$$

Let $(d^k, s^k) \in \mathbb{R}^{n+l}$ be a solution of this QP and $\lambda^k \equiv (\lambda_g^k, \lambda_h^k, \lambda_s^k) \in \mathbb{R}^{2l+m}$ be its corresponding KKT multipliers vector.

Step 2. (Termination check) If some stopping rule is satisfied, terminate. Otherwise, go to Step 3.
Step 3. (Penalty update) Let

$$\tilde{\rho}_k = \begin{cases} \rho_{k-1}, & \text{if } \rho_{k-1} > \max_{1\le i\le 2l+m} |\lambda_i^k|; \\ \delta_1 + \max_{1\le i\le 2l+m} |\lambda_i^k|, & \text{otherwise}, \end{cases} \tag{2.3}$$

$$\rho_k = \begin{cases} \tilde{\rho}_k, & \text{if } \sum_{i=1}^{l} s_i = 0; \\ \delta_2 + \tilde{\rho}_k, & \text{otherwise}, \end{cases} \tag{2.4}$$

$$\rho_k^g = \begin{cases} \tilde{\rho}_k, & \text{if } \sum_{i=1}^{l} s_i = 0; \\ \rho_{k-1}, & \text{otherwise}, \end{cases} \tag{2.5}$$

and

$$\rho_k^h = \begin{cases} \tilde{\rho}_k, & \text{if } r(x) = h(x); \\ \dfrac{\min\{(\lambda_h^k))^T (r(x) - h(x)), 0\}}{-2\|(r(x) - h(x))\|_2^2} + \delta_3, & \text{otherwise}. \end{cases} \tag{2.6}$$

Step 4. (Line search) Let $t_k = \tau^{i_k}$, where i_k is the smallest nonnegative integer i which satisfies the following inequality:

$$\theta_{\rho_k^g,\rho_k^h}(x^k + \tau^i d^k) \le \theta_{\rho_k^g,\rho_k^h}(x^k) - \sigma\tau^i (d^k)^T B_k d^k. \tag{2.7}$$

Step 5. (Update) Let $x^{k+1} = x^k + t_k d^k$. Choose a symmetric positive definite matrix $B_{k+1} \in \mathbb{R}^{n\times n}$. Set $k := k+1$. Go to Step 1.

It is well-known that the direction search and the step length determination are two critical steps amongst all SQP methods or its variants. In the direction search step of our algorithm, we further improve the prospect of feasibility of the QP subproblem by solving a linear least square problem (2.1), compared with the modified SQP method in [3]. This idea directly comes from [1]. However, our algorithm, including penalty parameter update and step length determination, is very different from [1].

Since B_k for each k is a symmetric positive definite, and QP (2.2) is always feasible with some vector $s \in \mathbb{R}^l$ sufficiently large, the search direction and the corresponding multipliers vector are also well-defined. The following lemma is useful in proving that d^k is a descent direction of the merit function.

Lemma 2.1. *If $d^k \neq 0$ for each k, where d^k is a solution QP (2.2) with $x = x^k$, then we have*

$$(\|h(x)\|_2^2)'(x^k; d^k) = -2\|h'(x^k)d^k\|_2^2 \le 0. \tag{2.8}$$

Proof. Since d^k satisfies $h'(x^k)d^k + h(x^k) = r(x^k)$, it must solve the least square problem (2.1). Therefore, it is a solution of the following linear equation:

$$h'(x)^T h'(x) d^k = -h'(x)^T h(x). \tag{2.9}$$

From (2.9), we have

$$\begin{aligned}(\|h(x)\|_2^2)'(x^k;d^k) &= 2(d^k)^T h'(x^k)^T h(x^k)\\ &= -2(d^k)^T h'(x)^T h'(x)d^k\\ &= -2\|h'(x^k)d^k\|_2^2\\ &\le 0.\end{aligned}$$

□

It is easy to see that $r(x^k) = h(x^k)$ if and only if the equality holds in (2.8).

The next lemma states that for $d^k = 0$, x^k turns out to be the critical point of the merit function under some condition.

Lemma 2.2. *Let $d^k = 0$ be a solution QP (2.2) with $x = x^k$. If $h'(x^k)^T\lambda_h^k = 0$, then x^k is a critical point of $\theta_{\rho_k^g,\rho_k^h}$ with ρ_k^g, ρ_k^h being defined as (2.5) and (2.6), respectively.*

Proof. Since $d^k = 0$ is a solution QP (2.2) with $x = x^k$, there must exist a multiplier vector $\lambda^k = (\lambda_g^k, \lambda_h^k, \lambda_s^k) \in \mathbb{R}^{2l+m}$ such that the following KKT conditions hold:

$$\begin{aligned}&\nabla f(x^k) - g'(x^k)\lambda_g^k + h'(x^k)\lambda_h^k = 0,\\ &\rho_{k-1}e - \lambda_g^k - \lambda_s^k = 0,\\ &g(x^k) \ge -s^k, \lambda_g^k \ge 0, (\lambda_g^k)^T(g(x^k)+s^k) = 0,\\ &s^k \ge 0, \lambda_s^k \ge 0, (\lambda_s^k)^T s = 0,\\ &h(x^k) = r(x^k).\end{aligned} \tag{2.10}$$

Recall that x^k is a critical point of $\theta_{\rho_k^g,\rho_k^h}$ and is equivalent to

$$\theta'_{\rho_k^g,\rho_k^h}(x^k;d) \ge 0, \quad \forall d \in \mathbb{R}^n.$$

To prove the required results, we need the following two inequalities:

$$\rho_k^g\left[\sum_{g_i(x^k)<0} -g_i'(x^k)d + \sum_{g_i(x^k)=0} \max(-g_i'(x^k)d, 0) + \sum_{g_i(x^k)>0} 0\right] \ge -(\lambda_g^k)^T g'(x^k)d; \tag{2.11}$$

$$2\rho_k^h h(x^k)^T h'(x^k)d \ge (\lambda_h^k)^T h'(x^k)d. \tag{2.12}$$

First, we prove the inequality (2.11). In the case that $g_i(x^k) < 0$, we have $s_i^k > 0$ and $(\lambda_s^k)_i = 0$, hence $\rho_{k-1} = (\lambda_g^k)_i$ from KKT conditions (2.10). Since for this case, $\sum s_i^k \neq 0$, we have $\rho_k^g = \rho_{k-1} = (\lambda_g^k)_i$ from (2.5). Therefore,

$$\rho_k^g \sum_{g_i(x^k)>0} -g_i'(x^k)d = -\sum_{g_i(x^k)<0} (\lambda_g^k)_i g_i'(x^k)d.$$

In the case that $g_i = 0$, if $g'(x^k)d < 0$, hence $\max(-g_i'(x^k)d, 0) = -g_i'(x^k)d$, then we have

$$\rho_k^g \sum_{g_i(x^k)=0} \max(-g_i'(x^k)d, 0) \ge -\sum_{g_i(x^k)=0} (\lambda_g^k)_i g_i'(x^k)d.$$

Otherwise, $\max(-g_i'(x^k)d, 0) = 0 \ge (\lambda_g^k)_i g_i'(x^k)d$.

In the case that $g_i(x^k) > 0$, since $s_i^k \ge 0$, hence $g_i(x^k) + s_i^k > 0$ and we have $(\lambda_g^k)_i = 0$.

From the above argument, we can deduce that inequality (2.11) holds.

The second inequality (2.12) can be proved by using condition $h^{'}(x^k)^T\lambda_h^k = 0$ and

$$h^{'}(x^k)^T h(x^k) = h^{'}(x^k)^T h^{'}(x^k) d^k = 0.$$

Moreover, from it we have that the equality holds in (2.12).

By the inequalities (2.11) and (2.12), it follows from the first equality in the KKT conditions (2.10) that for all $d \in \mathbb{R}^n$,

$$\theta^{'}_{\rho_k^g,\rho_k^h}(x^k; d) \geq (\nabla f(x^k) - g^{'}(x^k)\lambda_g^k + h^{'}(x^k)\lambda_h^k)^T d = 0.$$

□

Remark 2.3. The condition $h^{'}(x^k)^T\lambda_h^k = 0$ actually requires that the vector λ_h^k belongs to the null space of the matrix $\nabla h(x^k)$.

The last lemma in this section states that for every $d^k \neq 0$, it must be the descent direction of the merit function, which is important in making sure that the proposed algorithm is valid, in particular, the line search step can be finished in a finite number of times.

Lemma 2.4. *Let (d^k, s^k) be a solution of QP (2.2), and ρ_k^g, ρ_k^h be defined as in (2.5) and (2.6), respectively. Suppose that $d^k \neq 0$, then*

$$\begin{aligned}\theta^{'}_{\rho_k^g,\rho_k^h}(x^k; d^k) &\leq (\nabla f(x^k)^T d^k - (\lambda_g^k)^T g^{'}(x^k) d^k + (\lambda_h^k)^T h^{'}(x^k) d^k \\ &\leq -(d^k)^T B_k d^k < 0.\end{aligned} \tag{2.13}$$

Proof. Since (d^k, s^k) is a solution of QP (2.2) with $x = x^k$, there must exist a multiplier vector $\lambda^k = (\lambda_g^k, \lambda_h^k, \lambda_s^k) \in \mathbb{R}^{2l+m}$ such that the following KKT conditions hold:

$$\nabla f(x^k) + B_k d^k - g^{'}(x^k)\lambda_g^k + h^{'}(x^k)\lambda_h^k = 0, \tag{2.14}$$

$$\rho_{k-1} e - \lambda_g^k - \lambda_s^k = 0, \tag{2.15}$$

$$g^{'}(x^k) d^k + g(x^k) \geq -s^k, \lambda_g^k \geq 0, \lambda_g^k)^T (g^{'}(x^k) d^k + g(x^k) + s^k) = 0, \tag{2.16}$$

$$s^k \geq 0, \lambda_s^k \geq 0, (\lambda_s^k)^T s = 0, \tag{2.17}$$

$$h^{'}(x^k) d^k + h(x^k) = r(x^k). \tag{2.18}$$

Recall that

$$\begin{aligned}\theta^{'}_{\rho_k^g,\rho_k^h}(x^k; d^k) = \nabla f(x^k)^T d^k + \rho_k^g \Bigg[&\sum_{g_i(x^k)<0} -g_i^{'}(x^k) d^k \\ &+ \sum_{g_i(x^k)=0} \max(-g_i^{'}(x^k) d^k, 0) + \sum_{g_i(x^k)>0} 0 \Bigg] - 2\rho_k^h \|h^{'}(x^k) d^k\|_2^2.\end{aligned} \tag{2.19}$$

We first prove that the following inequalities hold:

$$\rho_k^g \left[\sum_{g_i(x^k)<0} -g_i^{'}(x^k) d^k + \sum_{g_i(x^k)=0} \max(-g_i^{'}(x^k) d^k, 0) + \sum_{g_i(x^k)>0} 0 \right] \leq -(\lambda_g^k)^T g^{'}(x^k) d^k; \tag{2.20}$$

$$-2\rho_k^h\|h^{'}(x^k)d^k\|_2^2 \le (\lambda_h^k)^T h^{'}(x^k)d^k. \tag{2.21}$$

It is easy to prove inequality (2.20) by using Lemma 2.1 and the penalty update rule (2.6). Here, we only prove (2.20).

In the case that $\sum_{i=1}^l s_i^k = 0$, we have $s_i^k = 0$ for each $i \in \{1, 2, \ldots, l\}$. If $g_i < 0$, then $-\nabla g_i(x^k)^T d^k \le g_i(x^k) + s_i^k = g_i < 0$, it follows from $\rho_k^g = \tilde{\rho}_k \ge (\lambda_g^k)_i$ that

$$-\rho_k^g \nabla g_i(x^k)^T d^k \le -(\lambda_g^k)_i \nabla g_i(x^k)^T d^k.$$

If $g_i = 0$, then $-\nabla g_i(x^k)^T d^k \le g_i(x^k) + s_i^k = 0$, hence $\max(-\nabla g_i(x^k)^T d^k, 0) = 0$, and

$$-(\lambda_g^k)_i \nabla g_i(x^k)^T d^k = -(\lambda_g^k)_i (g_i(x^k) + s_i^k) = 0.$$

If $g_i > 0$, then $-(\lambda_g^k)_i \nabla g_i(x^k)^T d^k = -(\lambda_g^k)_i (g_i(x^k) + s_i^k) > 0$.

In the case that $\sum_{i=1}^l s_i^k \ne 0$, we have $\rho_k^g = \rho_{k-1}$. If $s_i^k > 0$, then from (2.17) we have $(\lambda_s^k)_i = 0$, hence from (2.15) we have $\rho_{k-1} = (\lambda_g^k)_i$. It directly follows that

$$-\sum_{g_i(x^k)<0} (\lambda_g^k)_i \nabla g_i(x^k)^T d^k = -\sum_{g_i(x^k)<0} \rho_{k-1} \nabla g_i(x^k)^T d^k = -\sum_{g_i(x^k)<0} \rho_k^g \nabla g_i(x^k)^T d^k,$$

$$\begin{aligned}\sum_{g_i(x^k)=0} \rho_k^g \max(-\nabla g_i(x^k)^T d^k, 0) &= \sum_{g_i(x^k)=0} \max(-(\lambda_g^k)_i \nabla g_i(x^k)^T d^k, 0) \\ &= \sum_{g_i(x^k)=0} \max((\lambda_g^k)_i s_i^k, 0) \\ &= -\sum_{g_i(x^k)=0} (\lambda_g^k)_i \nabla g_i(x^k)^T d^k.\end{aligned}$$

For $g_i > 0$, we also have

$$-(\lambda_g^k)_i \nabla g_i(x^k)^T d^k = (\lambda_g^k)_i (g_i(x^k) + s_i^k) > 0.$$

If $s_i^k = 0$, then $\nabla g_i(x^k)^T d^k + g_i(x^k) \ge 0$, $(\lambda_g^k)_i \ge 0$ and $(\lambda_g^k)_i(\nabla g_i(x^k) d^k + g_i(x^k)) = 0$. Therefore,

$$\begin{aligned} &-\nabla g_i(x^k)^T d^k < 0, && \text{for } g_i(x^k) < 0; \\ &\nabla g_i(x^k)^T d^k \ge 0, && \text{for } g_i(x^k) = 0; \\ &-(\lambda_g^k)_i \nabla g_i(x^k)^T d^k = (\lambda_g^k)_i g_i(x^k) \ge 0, && \text{for } g_i(x^k) > 0. \end{aligned}$$

Since from (2.15),we have $\rho_{k-1} = (\lambda_g^k)_i + (\lambda_s^k)_i \ge (\lambda_g^k)_i$, and combined with above argument, we can obtain the inequality (2.20).

Multiply (2.14) by d^k, and from (2.20) and (2.21), we obtain that

$$\begin{aligned}\theta^{'}_{\rho_k^g, \rho_k^h}(x^k; d^k) &\le (\nabla f(x^k)^T d^k - (\lambda_g^k)^T g^{'}(x^k) d^k + (\lambda_h^k)^T h^{'}(x^k) d^k \\ &\le -(d^k)^T B_k d^k < 0.\end{aligned}$$

□

From the above argument, we know that our modified SQP method is well-defined.

3. GLOBAL CONVERGENCE

In this section, we study the global convergence of the algorithm. For this, we assume that $d^k \neq 0$ for each k, and let $\{x^k\}$ be a infinite iterate sequence generated by the algorithm. Moreover, we make the following blanket assumptions:

(A_1). For all k, there exist two positive constants $\alpha < \beta$ satisfying

$$\alpha\|d\|^2 \leq d^T B_k d \leq \beta\|d\|^2, \quad \forall d \in \mathbb{R}^n.$$

(A_2). After finitely many iterations, $\rho_k \equiv \rho_1^*$, $\rho_k^h \equiv \rho_2^*$.

Lemma 3.1. *Under (A_1) and (A_2), suppose that x^* is a cluster point of $\{x^k\}$, i.e., for some subset κ, $\lim_{k(\in\kappa)\to\infty} x^k = x^*$, then the following conclusions hold.*

1. $\sum_{i=1}^l s_i^k = 0,$ *for $k \in \kappa$ large enough;*
2. *The multiplier sequences $\{\lambda_g^k\}_{k\in\kappa}$,$\{\lambda_h^k\}_{k\in\kappa}$, $\{\lambda_s^k\}_{k\in\kappa}$ and the penalty parameter sequence $\{\rho_k^g\}_{k\in\kappa}$ are bounded;*
3. *The direction sequence $\{d^k\}_{k\in\kappa}$ is bounded;*
4. *If $\lim_{k(\in\kappa)\to\infty} d^k = d^*$, $\lim_{k(\in\kappa)\to\infty} B_k = B_*$, $\lim_{k(\in\kappa)\to\infty} \rho_k^g = \rho_*^g$, $\lim_{k(\in\kappa)\to\infty} \rho_k = \rho_*$, $\lim_{k(\in\kappa)\to\infty} \rho_k^h = \rho_*^h$, then $(d^*, 0)$ is the solution of the (2.2) with $x = x^*$, $r(x) = 0$. Moreover, the following inequality holds:*

$$\theta'_{\rho_*^g,\rho_*^h}(x^*; d^*) \leq -(d^*)^T B_* d^*. \tag{3.1}$$

5. *If $d^* = 0$ and $r(x^*) = 0$, then x^* is feasible for the primary problem, and is a feasible stationary point.*

Proof. Since for all k large enough, $\rho_k \equiv \rho_1^*$, we can deduce that $\sum_{i=1}^l s_i^k = 0$ after finite steps by the penalty update rule (2.4), hence from the boundedness of $\{\rho_k\}$ we know that $\{\tilde{\rho_k}\}$ is also bounded. By the penalty update rule (2.3), it follows that after finite steps, we have

$$\rho_{k-1} \geq \max_{1\leq i\leq 2l+m} |\lambda_i^k|.$$

By assumption (A_2), we obtain that for $k \in \kappa$, $\{\lambda_g^k\}$, $\{\lambda_h^k\}$ and $\{\lambda_s^k\}$ are bounded.

Since $\sum_{i=1}^l s_i^k = 0$ after finite steps, so for $k(\in \kappa)$ large enough we have that

$$\rho_k^g \equiv \tilde{\rho}_k,$$

hence that $\{\rho_k^g\}$ is bounded. The first and second conclusions above have been proved.

Next, we prove that the third conclusion holds.

From the boundedness of $\{\lambda_g^k\}$ and $\{\lambda_h^k\}$, without loss of generality, we assume that

$$\lim_{k(\in\kappa)\to\infty} \lambda_g^k \equiv \lambda_g^*, \quad \lim_{k(\in\kappa)\to\infty} \lambda_h^k \equiv \lambda_h^*.$$

Using the KKT conditions (2.14), we obtain that

$$\lim_{k(\in\kappa)\to\infty} B_k d^k = -\nabla f(x^*) + g'(x^*)\lambda_g^* - h'(x^*)\lambda_h^*.$$

Thus, we can deduce that $\{B_k d^k : k \in \kappa\}$ is bounded. By assumption (A_1), we have

$$\|d^k\|^2 \leq \frac{1}{\alpha}\|d^k\|\|B_k d^k\| \leq \frac{1}{\alpha}\|d^k\|M, \ \ \forall k \in \kappa, \tag{3.2}$$

where M is a constant scalar large enough.

If we assume that $\|d^k\| \neq 0$, then from (3.2) we know that $\|d^k\| \leq \frac{1}{2}M$. i.e. the direction sequence $\{d^k : k \in \kappa\}$ is bounded.

Then, we prove the forth conclusion.

By the second and the third conclusions, we can assume that

$$\lim_{k(\in\kappa)\to\infty} d^k = d^*, \qquad \lim_{k(\in\kappa)\to\infty} B_k = B_*,$$
$$\lim_{k(\in\kappa)\to\infty} \rho_k^g = \rho_*^g, \qquad \lim_{k(\in\kappa)\to\infty} \rho_k^h = \rho_2^*,$$
$$\lim_{k(\in\kappa)\to\infty} \rho_k^s = \rho_*^s, \qquad \lim_{k(\in\kappa)\to\infty} \rho_k = \rho_1^*,$$

then from the KKT conditions (2.14) – (2.18) we have

$$\begin{gathered} \nabla f(x^*) + B_* d^* - g'(x^*)\lambda_g^* + h'(x^*)\lambda_h^* = 0, \\ \rho_1^* e - \lambda_g^* - \lambda_s^* = 0, \\ g'(x^*)d^* + g(x^*) \geq -s^*, \lambda_g^* \geq 0, \\ (\lambda_g^*)^T(g'(x^*)d^* + g(x^*) + s^*) = 0, \\ s^* \geq 0, \lambda_s^* \geq 0, (\lambda_s^*)^T s = 0, \\ h'(x^*)d^* + h(x^*) = r(x^*). \end{gathered} \tag{3.3}$$

It shows that $(d^*, 0)$ is a solution of the following problem:

$$\begin{aligned} \min \quad & \nabla f(x^*)^T d + \tfrac{1}{2} d^T B_* d + \rho_1^* \textstyle\sum_{i=1}^l s_i \\ \text{s.t.} \quad & g'(x^*)d + g(x^*) \geq -s, \\ & h'(x^*)d + h(x^*) = r(x^*), \\ & s \geq 0. \end{aligned} \tag{3.4}$$

Therefore, d^* is a solution of the following problem:

$$\begin{aligned} \min \quad & \nabla f(x^*)^T d + \tfrac{1}{2} d^T B_* d \\ \text{s.t.} \quad & g'(x^*)d + g(x^*) \geq 0, \\ & h'(x^*)d + h(x^*) = r(x^*). \end{aligned} \tag{3.5}$$

Also since $\lim_{k(\in\kappa)\to\infty} \rho_k^h = \rho_*^h$, so from the penalty update rule (2.6), we obtain that $r(x^k) = h(x^k)$ after finite steps. Hence we have $r(x^*) = h(x^*)$ in the problem (3.5). The inequality (3.1) can be easily proved under assumption (A_2).

In what follows, we prove the last conclusion.

If $d^* = 0$ and $r(x^*) = 0$, then from $r(x^*) = h(x^*)$ proved above we can obtain that $g(x^*) \geq 0$, $h(x^*) = 0$. i.e. x^* is feasible for the primary problem. Also since $d^* = 0$ is a solution of the QP subproblem (1.2), we can deduce that x^* is a feasible stationary point of the original problem. □

The following lemma can be proved similar to the corresponding result in [3].

Lemma 3.2. *Supposing that*

$$\lim_{k(\in\kappa)\to\infty} x^k = x^*, \qquad \lim_{k(\in\kappa)\to\infty} t_k = 0, \qquad \lim_{k(\in\kappa)\to\infty} \rho_k^g = \rho_*^g,$$
$$\lim_{k(\in\kappa)\to\infty} \rho_k^h = \rho_2^*, \qquad \lim_{k(\in\kappa)\to\infty} d^k = d^*,$$

then we have

$$\limsup_{k(\in\kappa)\to\infty} \frac{\theta_{\rho_*^g,\rho_2^*}(x^k + t_k d^k) - \theta_{\rho_*^g,\rho_2^*}(x^k)}{t_k} \leq \theta'_{\rho_*^g,\rho_2^*}(x^*; d^*).$$

Theorem 3.3. *Suppose that* $\lim_{k(\in\kappa)\to\infty} x^k = x^*$. *Under assumptions* (A_1) *and* (A_2), *we have* $\lim_{k(\in\kappa)\to\infty} d^k = d^* = 0$. *Therefore,* x^* *is a generalized stationary point of the primary problem (1.1). If* $r(x^*) = 0$, *then* x^* *is a feasible stationary point.*

Proof. From assumption (A_2) and Lemma 3.1, we know that for k large enough, the following equality holds:

$$\theta_{\rho_k^g,\rho_k^h}(x^k) \equiv \theta_{\rho_*^g,\rho_2^*}(x^k).$$

By Lemma 2.4, we know that $\{\theta_{\rho_*^g,\rho_2^*}(x^k) : k \in \kappa\}$ is a monotonically decreasing sequence and lower bounded, hence from Lemma 3.1, we have the following limits (if necessary, we can choose some subsequence):

$$\lim_{k(\in\kappa)\to\infty} d^k = d^*, \quad \lim_{k(\in\kappa)\to\infty} B_k = B_*.$$

Next, we prove that $d^* = 0$.

If the cluster point t^* of step length sequence $\{t_k : k \in \kappa\}$ is nonzero, then from the line search step of the proposed algorithm, we have

$$\lim_{k(\in\kappa)\to\infty} t_k (d^k)^T B_k d^k = 0,$$

or

$$t^*(d^*)^T B_* d^* = 0,$$

hence by the positive definiteness of B_*, we can deduce that $d^* = 0$.

If $t^* = 0$, then

$$\theta_{\rho_*^g,\rho_2^*}\left(x^k + \frac{t_k}{\tau} d^k\right) - \theta_{\rho_*^g,\rho_2^*}(x^k) > -\sigma \frac{t_k}{\tau}(d^k)^T B_k d^k,$$

or

$$\begin{aligned} -\sigma \frac{t_k}{\tau}(d^k)^T B_k d^k &< \frac{\theta_{\rho_*^g,\rho_2^*}(x^k + \frac{t_k}{\tau} d^k) - \theta_{\rho_*^g,\rho_2^*}(x^k)}{\frac{t_k}{\tau}} \\ &\le \lim_{k(\in\kappa)\to\infty} \sup \frac{\theta_{\rho_*^g,\rho_2^*}(x^k + \frac{t_k}{\tau} d^k) - \theta_{\rho_*^g,\rho_2^*}(x^k)}{\frac{t_k}{\tau}} \\ &\le \theta'_{\rho_*^g,\rho_2^*}(x^*; d^*) \\ &\le -(d^*)^T B_* d^*, \end{aligned}$$

i.e. $(1-\sigma)(d^*)^T B_* d^* \le 0$. So for $\sigma \in (0,1)$, we have $(d^*)^T B_* d^* \le 0$. By the positive definiteness of B_*, we also obtain that $d^* = 0$.

At last, from the forth and fifth conclusion in Lemma 3.1, we can prove the desired results. (see [3, Proposition A.4]). □

REFERENCES

[1] J.V. BURKE AND S.-P. HAN, A robust sequential quadatic programming method, *Math. Programming.*, **43** (1989), 277–303.

[2] F.H. CLARKE, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, (1983).

[3] H.Y. JIANG AND D. RALPH, Smooth SQP methods for mathematical programs with nonlinear complementarity constraints, *SIAM J. Optim.*, **10** (2000), 779–808.