A globally and superlinearly convergent trust-region SQP method without a penalty function for nonlinearly constrained optimization

Hiroshi Yamashita* and Hiroshi Yabe†

September 2003 (revised July 2007)

Abstract

In this paper, we propose a new trust-region SQP method, which uses no penalty function, for solving nonlinearly constrained optimization problem. Our method consists of alternate two algorithms. Specifically, we alternately proceed the feasibility restoration algorithm and the objective function minimization algorithm. The global and superlinear convergence property of the proposed method is shown.

1 Introduction

In this paper, we consider the nonlinearly constrained optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) = 0, \qquad x \geq 0, \end{array} \tag{1}$$

where $f : \mathbf{R}^n \to \mathbf{R}$, $g_i : \mathbf{R}^n \to \mathbf{R}, i = 1, \ldots, m$ are smooth functions and $g(x) = (g_1(x), \ldots, g_m(x))^T$. There are several numerical methods for solving the above problem, which include the augmented Lagrangian function method, the sequential quadratic programming (SQP) method, the trust-region SQP method, and the interior point method.

Most of the methods use penalty functions as merit functions in the line search strategy or the trust-region strategy. The penalty function is a linear combination of the objective function and some measure of the constraint violation, in which the objective function minimization and the constraints satisfaction are treated together within the framework of a single penalty function minimization problem. There are several difficulties associated with the use of penalty functions. A typical difficulty arises in choosing a penalty parameter.

*Mathematical Systems, Inc., 2-4-3, Shinjuku, Shinjuku-ku, Tokyo, Japan. hy@msi.co.jp

†Department of Mathematical Information Science, Faculty of Science, Tokyo University of Science, 1-3, Kagurazaka, Shinjuku-ku, Tokyo, Japan. yabe@rs.kagu.tus.ac.jp

In order to remedy these defects, some researchers proposed methods that use no penalty functions, in which the above two goals are treated separately. These methods are contained in the category of a multi-objective optimization. For example, Yamashita [5] proposed a method based on the SQP method for solving equality constrained optimization. In his method, the restoration algorithm that improved feasibility and the minimization algorithm that reduced the objective function value were done at each iteration. Fletcher and Leyffer [3] proposed the filter method for solving general constrained optimization problems, and they presented some numerical experiments that suggested the good performance of their method. Furthermore, Fletcher et al. [2] proved the global convergence property of the filter method within the framework of the trust-region SQP method. On the other hand, Ulbrich and Ulbrich [4] proposed non-monotone trust-region methods for nonlinear equality constrained optimization without a penalty function.

In this paper, we will propose a new method without a penalty function. The proposed method is based on the trust-region SQP method and consists of alternate two algorithms. Specifically, we will alternately proceed the feasibility restoration algorithm and the objective function minimization algorithm. The global and superlinear convergence property of the proposed method will be shown. The present research is motivated by the earlier paper by Yamashita [5]. This paper is organized as follows. In Section 2, we will give some notations used in the subsequent sections. In Section 3, we will describe the algorithms of our method, and in Section 4, we will prove the global convergence properties of the proposed method. In Section 5, the rate of convergence of our method will be analyzed.

Throughout this paper, $\|\cdot\|$ denotes the l_2 norm for vectors and the induced norm for matrices.

2 Notations

Let the Lagrangian function of problem (1) be defined by

$$L(w) = f(x) - y^t g(x) - z^t x, \tag{2}$$

where $w = (x, y, z)^t$, and $y \in \mathbf{R}^m$ and $z \in \mathbf{R}^n$ are the Lagrange multiplier vectors which correspond to the equality and inequality constraints, respectively. Then Karush-Kuhn-Tucker (KKT) conditions for optimality of the above problem are given by

$$r(w) = \begin{pmatrix} \nabla_x L(w) \\ g(x) \\ XZe \end{pmatrix} = 0, \quad x \geq 0,\ z \geq 0, \tag{3}$$

where

$$\nabla_x L(w) \quad = \quad \nabla f(x) - A(x)^t y - z,$$

$$
\begin{aligned}
A(x) &= \begin{pmatrix} \nabla g_1(x)^t \\ \vdots \\ \nabla g_m(x)^t \end{pmatrix}, \\
X &= \operatorname{diag}(x_1, \cdots, x_n), \\
Z &= \operatorname{diag}(z_1, \cdots, z_n), \\
e &= (1, \cdots, 1)^t \in \mathbf{R}^n.
\end{aligned}
$$

We define a first order approximation $f_l(x; s) : \mathbf{R}^n \to \mathbf{R}$ and a second order approximation $f_q(x; s) : \mathbf{R}^n \to \mathbf{R}$ to the objective function $f(x)$ by

$$
\begin{aligned}
f_l(x; s) &= f(x) + \nabla f(x)^t s \\
f_q(x; s) &= f(x) + \nabla f(x)^t s + \frac{1}{2} s^t H s,
\end{aligned}
$$

respectively, where $s \in \mathbf{R}^n$ is a step and $H \in \mathbf{R}^{n \times n}$ is a suitable symmetric matrix given below. We also define the differences

$$
\begin{aligned}
\Delta f_l(x; s) &\equiv f_l(x; s) - f(x) = \nabla f(x)^t s, \\
\Delta f_q(x; s) &\equiv f_q(x; s) - f(x) = \nabla f(x)^t s + \frac{1}{2} s^t H s, \\
\Delta f(x; s) &\equiv f(x + s) - f(x).
\end{aligned}
$$

The differences $\Delta f_l(x; s)$ and $\Delta f_q(x; s)$ are used as objective functions in LP and QP subproblems given below, respectively. The differences $\Delta f(x; s)$ and $\Delta f_q(x; s)$ are used in adapting trust-region radii. They are usually called the actual reduction and the predicted reduction, respectively.

3 Algorithm of Our Method

In this section, we present the algorithm of our proposed method. Our method aims to obtain a point that satisfies the KKT conditions. The main algorithm consists of the feasibility restoration algorithm and the objective function minimization algorithm. In both algorithms, the nonnegativity of primal and dual variables is maintained. The feasibility restoration algorithm finds a point that approximately satisfies the equality constraints within given tolerances. The objective function minimization algorithm decreases the objective function values, while the magnitude of the equality constraints is restricted within the tolerance given by the preceding restoration algorithm. In what follows, $\|r(w)\|_*$ is defined by the form

$$
\|r(w)\|_* = \max(\|\nabla_x L(w)\|, \ \|g(x)\|, \ \|XZe\|).
$$

First, we describe the main algorithm of our method.

[Main Algorithm of our method]

(Step 0) Initialization:

w_0 (with $x_0 \geq 0$), $\varepsilon > 0, \tau \in (0,1)$ are given. Set $k = 0$.

(Step 1) Set $\delta_k = \tau \|r(w_k)\|_*$.

(Step 2) Perform **Restoration Algorithm** by starting from the point x_k and find

$$w_{k+\frac{1}{2}} = (x_{k+\frac{1}{2}}, y_{k+\frac{1}{2}}, z_{k+\frac{1}{2}})^t$$

that satisfies

$$\|g(x_{k+\frac{1}{2}})\| < \delta_k, \quad x_{k+\frac{1}{2}} \geq 0 \quad \text{and} \quad z_{k+\frac{1}{2}} \geq 0.$$

If $w_{k+\frac{1}{2}}$ satisfies $\|r(w_{k+\frac{1}{2}})\|_* \leq \delta_k$, then set $w_{k+1} = w_{k+\frac{1}{2}}$ and go to Step 4.

(Step 3) By **Minimization Algorithm** with an initial point $x_{k+\frac{1}{2}} \geq 0$, find

$$w_{k+1} = (x_{k+1}, y_{k+1}, z_{k+1})^t$$

that satisfies

$$\|r(w_{k+1})\|_* \leq \delta_k, \quad x_{k+1} \geq 0 \quad \text{and} \quad z_{k+1} \geq 0.$$

(Step 4) Convergence Check:

If $\|r(w_{k+1})\|_* \leq \varepsilon$, then stop. Otherwise, set $k := k+1$ and go to Step 1. □

Next, we give the feasibility restoration algorithm, in which one QP subproblem is solved. The algorithm proceeds until the magnitude of the equality constraints becomes less than a prescribed tolerance δ.

[Restoration Algorithm]

(Step 0) Initialization:

w_0 (with $x_0 \geq 0$), $\delta > 0, \Delta_0 > 0, \varepsilon_0 \in (0,1)$ and $\beta \in (0,1)$ are given. Set $k = 0$.

(Step 1) Calculate a matrix G_k. (G_k is the Hessian matrix $\nabla_x^2 L(w_k)$ or its approximation.)

(Step 2) QP subproblem:

Calculate the step s_k and multipliers $(y_{k+1}, z_{k+1})^t$ corresponding to the equality and nonnegativity constraints by solving the QP subproblem

$(QP_R(x_k))$

$$\begin{array}{ll} \text{minimize} & \dfrac{1}{2} s^t G_k s + \nabla f(x_k)^t s \\ \text{subject to} & g(x_k) + A(x_k)s = 0, \quad x_k + s \geq 0, \quad \|s\|_\infty \leq \Delta_k. \end{array}$$

If necessary, the trust-region radius Δ_k is adapted so that the QP subproblem is feasible.

(Step 3) Line Search:

Find the least nonnegative integer l_k that satisfies

$$\|g(x_k + \beta^{l_k} s_k)\| < \max\{\delta, (1 - \varepsilon_0 \beta^{l_k})\|g(x_k)\|\},$$

and set $\alpha_k = \beta^{l_k}$.

(Step 4) Set $x_{k+1} = x_k + \alpha_k s_k$ and set $w_{k+1} = (x_{k+1}, y_{k+1}, z_{k+1})^t$.

(Step 5) Convergence Check:

If $\|g(x_{k+1})\| < \delta$, then stop.

(Step 6) Set $k := k + 1$ and go to Step 1. □

The given constant δ in Step 0 means a tolerance δ_k used in Step 1 of [Main Algorithm]. In Step 2, the actual form of the trust-region constraint $\|s\|_\infty \leq \Delta_k$ is defined by the bound constraints $-\Delta_k e \leq s \leq \Delta_k e$, where $e = (1, ..., 1)^t$.

Remark: Unlike usual trust-region methods, our trust-region strategy does not make the radius tend to zero. The trust-region radius is uniformly bounded below.

To close this section, we give the minimization algorithm, in which one LP subproblem and two QP subproblems are solved. The LP subproblem plays an important role in calculating estimates of the multipliers y and z, while the QP subproblems find a search direction of the primal variable x. The algorithm proceeds until an approximate KKT point is obtained within a prescribed tolerance δ.

[Minimization Algorithm]

(Step 0) Initialization:

w_0, $\delta > 0, \Delta_0 > 0$, $\Delta_{T_0} > 0$ and $\beta \in (0, 1)$ are given, where $\Delta_{T_0} \leq \Delta_0$, and x_0 satisfies $x_0 \geq 0$ and $\|g(x_0)\| < \delta$. Set $k = 0$.

(Step 1) LP subproblem:

Calculate the step d_k and multipliers $(y_{k+1}, z_{k+1})^t$ corresponding to the equality and nonnegativity constraints by solving the LP problem

$(LP(x_k))$

$$\begin{array}{rl} \text{minimize} & \Delta f_l(x_k; d) = \nabla f(x_k)^t d \\ \text{subject to} & A(x_k)d = 0, \quad x_k + d \geq 0, \quad \|d\|_\infty \leq 1. \end{array}$$

Set $w_k = (x_k, y_{k+1}, z_{k+1})^t$.

(Step 2) Convergence Check:

If w_k satisfies $\|r(w_k)\| \le \delta$, then stop.

(Step 3) Two QP subproblems:

Calculate a symmetric matrix $H_k \in \mathbf{R}^{n\times n}$. Obtain two steps s_{T_k} and s_k by solving the following two QP subproblems:

$(QP_T(x_k))$

$$\begin{array}{ll} \text{minimize} & \Delta f_q(x_k; s_T) = \frac{1}{2} s_T^t H_k s_T + \nabla f(x_k)^t s_T \\ \text{subject to} & A(x_k) s_T = 0, \quad x_k + s_T \ge 0, \quad \|s_T\|_\infty \le \Delta_{T_k} \end{array}$$

and

$(QP(x_k))$

$$\begin{array}{ll} \text{minimize} & \Delta f_q(x_k; s) = \frac{1}{2} s^t H_k s + \nabla f(x_k)^t s \\ \text{subject to} & g(x_k) + A(x_k) s = 0, \quad x_k + s \ge 0, \quad \|s\|_\infty \le \Delta_k, \end{array}$$

where Δ_k is chosen such that $\Delta_{T_k} \le \Delta_k$ and subproblem $(QP(x_k))$ is feasible.

(Step 4) Set $\bar{s}_k = \left(\min \left\{ \frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1 \right\} \right) s_k$. Find the least nonnegative integer l_k that satisfies

$$\Delta f_q(x_k; (1-\beta^{l_k}) s_{T_k} + \beta^{l_k} \bar{s}_k) \le \frac{1}{2} \Delta f_q(x_k; s_{T_k}), \tag{4}$$

and set $\rho_k = \beta^{l_k}$. Set $s_{\rho_k} = (1-\rho_k) s_{T_k} + \rho_k \bar{s}_k$.

(Step 5) Trust Region Radius:

If $\|g(x_k + s_{\rho_k})\| \ge \delta$, then set $\Delta_{T_{k+1}} = \frac{1}{2}\Delta_{T_k}$.

Otherwise

$$\left\{ \begin{array}{l} \text{if } \Delta f(x_k; s_{\rho_k}) > \frac{1}{4} \Delta f_q(x_k; s_{\rho_k}), \text{then set } \Delta_{T_{k+1}} = \frac{1}{2}\Delta_{T_k} \\ \text{if } \Delta f(x_k; s_{\rho_k}) \le \frac{3}{4} \Delta f_q(x_k; s_{\rho_k}), \text{then set } \Delta_{T_{k+1}} = 2\Delta_{T_k} \\ \text{otherwise } \Delta_{T_{k+1}} = \Delta_{T_k}. \end{array} \right.$$

(Step 6) If $\Delta f(x_k; s_{\rho_k}) \le 0$ and $\|g(x_k + s_{\rho_k})\| < \delta$, then set $x_{k+1} = x_k + s_{\rho_k}$. Otherwise, set $x_{k+1} = x_k$.

(Step 7) Set $k := k+1$ and go to Step 1. □

The given constant δ in Step 0 means a tolerance δ_k used in Step 2 of **Main Algorithm**. In Step 1 and Step 3, the actual forms of the trust-region constraints $\|d\|_\infty \leq \Delta_k$, $\|s_T\|_\infty \leq \Delta_{T_k}$ and $\|s\|_\infty \leq \Delta_k$ are defined by the bound constraints $-\Delta_k e \leq d \leq \Delta_k e$, $-\Delta_{T_k} e \leq s_T \leq \Delta_{T_k} e$ and $-\Delta_k e \leq s \leq \Delta_k e$, respectively. In Step 3, several kinds of matrices can be considered as H_k. Typical choices are $\nabla^2 f(x_k)$, $\nabla_x^2 L(w_k)$ and their approximations. Note that subproblems $(LP(x_k))$ and $(QP_T(x_k))$ are always feasible, because the zero vector is a feasible solution. Thus the following relations hold

$$\Delta f_l(x_k; d_k) \leq \Delta f_l(x_k; 0) = 0,$$

and

$$\Delta f_q(x_k; s_{T_k}) \leq \Delta f_q(x_k; 0) = 0.$$

Roughly speaking, subproblems $(QP_T(x_k))$ and $(QP(x_k))$ produce a descent search direction for the objective function and the Newton direction for the nonlinear equations $g(x) = 0$, respectively. This is a reason why we combine the two directions in Step 4. By choosing a step s_{ρ_k}, which satisfies (4), as close to the Newton direction as possible, we can expect to have a fast convergence. Since the estimate

$$\|\bar{s}_k\|_\infty = \left(\min\left\{\frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\}\right)\|s_k\|_\infty \leq \|s_{T_k}\|_\infty \leq \Delta_{T_k}$$

holds, we have

$$\|s_{\rho_k}\|_\infty \leq (1-\rho_k)\|s_{T_k}\|_\infty + \rho_k\|\bar{s}_k\|_\infty \leq \Delta_{T_k}.$$

We note that the generated sequence $\{w_k\}$ satisfies the conditions $x_k \geq 0$, $z_k \geq 0$ and $\|g(x_k)\| < \delta$.

4 Global convergence

In this section, we devote to prove the global convergence of our algorithm. For this purpose, we make the following assumptions.

Assumption G

(G1) The functions f and $g_i, i = 1, ..., m$, are twice continuously differentiable.

(G2) The set $\{x \in \mathbf{R}^n \,|f(x) \leq f(x_0)\} \cap \{x \in \mathbf{R}^n|\ \|g(x)\| \leq \delta_0\} \cap \{x \in \mathbf{R}^n|\ x \geq 0\}$ is compact for any x_0, where the set $\{x \in \mathbf{R}^n \,|f(x) \leq f(x_0)\}$ denotes the level set of the objective function $f(x)$ at $x_0 \in \mathbf{R}^n$.

(G3) The matrices G_k and H_k are uniformly bounded.

(G4) There exists a positive constant $\bar{\Delta}$ such that

$$0 < \Delta_{T_k} \leq \Delta_k \leq \bar{\Delta} \quad \text{for all } k.$$

4.1 Global convergence of Restoration Algorithm

The following theorem implies the global convergence of Restoration Algorithm.

Theorem 1 *Suppose that Assumption G holds. Let the sequence* $\{x_k\}$ *be generated by Restoration Algorithm. Then the line search procedure in Step 3 terminates in a finite number of iterations. Furthermore, any accumulation point* x_∞ $(\in \mathbf{R}^n)$ *of the sequence* $\{x_k\}$ *satisfies* $g(x_\infty) = 0$.

Proof. Suppose that $g(x_k) \neq 0$ for any k. Letting

$$\Phi(x) = \|g(x)\|^2,$$

we have

$$\nabla\Phi(x) = 2A(x)^t g(x).$$

Since the search direction s_k obtained in Step 3 satisfies

$$\nabla\Phi(x_k)^t s_k = 2g(x_k)^t A(x_k) s_k = 2g(x_k)^t(-g(x_k)) = -2\Phi(x_k) < 0,$$

s_k is a descent direction for $\Phi(x)$ at x_k.

For any fixed number $\varepsilon_0 \in (0,1)$, there exists a positive number α such that

$$\begin{aligned}\Phi(x_k + \alpha s_k) &\leq \Phi(x_k) + \varepsilon_0 \alpha \nabla\Phi(x_k)^t s_k \\ &= (1 - 2\varepsilon_0\alpha)\Phi(x_k).\end{aligned}$$

By the facts that s_k is uniformly bounded for each k and that $\nabla\Phi(x)$ is uniformly continuous on the compact set, there exists an integer $L < \infty$ such that we can find an integer $l_k \in [0, L]$ satisfying

$$\Phi(x_k + \beta^{l_k} s_k) \leq (1 - 2\varepsilon_0\beta^{l_k})\Phi(x_k).$$

Thus we have

$$\begin{aligned}\|g(x_k + \beta^{l_k} s_k)\| &\leq \sqrt{1 - 2\varepsilon_0\beta^{l_k}}\|g(x_k)\| \\ &\leq (1 - \varepsilon_0\beta^{l_k})\|g(x_k)\|.\end{aligned}$$

The finite termination property is obvious from $l_k \in [0, L]$ for each k, which implies that Step 3 is well-defined.

Since the preceding equation implies

$$\|g(x_{k+1})\| \leq (1 - \varepsilon_0\beta^L)\|g(x_k)\|$$

at each iteration and $0 < 1 - \varepsilon_0\beta^L < 1$ is satisfied, the magnitude of $\|g(x_k)\|$ approaches zero. Therefore, the theorem is proved. □

4.2 Global convergence of Minimization Algorithm

In this section, we study the global convergence property of Minimization Algorithm. We begin with the following lemma, which suggests the roles of the differences $\Delta f_l(x_k; d_k)$ and $\Delta f_q(x_k; s_{T_k})$.

Lemma 1 *Suppose that Assumption G holds. Let d_k, y_{k+1} and z_{k+1} be an optimal solution and corresponding multipliers for problem LP(x_k), respectively. Then the following hold:*
(i) *If $\Delta f_l(x_k; d_k) = 0$ at some k, then the point $w_k = (x_k, y_{k+1}, z_{k+1})^t$ satisfies*

$$\|g(x_k)\| < \delta, \quad \nabla_x L(w_k) = 0, \quad X_k Z_{k+1} e = 0, \quad x_k \geq 0, \quad z_{k+1} \geq 0,$$

for a given $\delta > 0$.
(ii) *If there exists a subsequence $K \subset \{0, 1, 2, \ldots\}$ such that*

$$\lim_{k\to\infty,\ k\in K} \Delta f_l(x_k; d_k) = 0,$$

then any accumulation point $w_\infty = (x_\infty, y_\infty, z_\infty)^t$ of $\{(x_k, y_{k+1}, z_{k+1})^t\}$ satisfies

$$\|g(x_\infty)\| < \delta, \quad \nabla_x L(w_\infty) = 0, \quad X_\infty Z_\infty e = 0, \quad x_\infty \geq 0, \quad z_\infty \geq 0,$$

for a given $\delta > 0$.
(iii) *If $\Delta f_q(x_k; s_{T_k}) = 0$ at some k, then $\Delta f_l(x_k; d_k) = 0$ holds.*
(iv) *If there exists a subsequence $K \subset \{0, 1, 2, \ldots\}$ such that*

$$\lim_{k\to\infty,\ k\in K} \Delta f_q(x_k; s_{T_k}) = 0 \quad \text{and} \quad \liminf_{k\to\infty,\ k\in K} \Delta_{T_k} > 0,$$

then $\lim_{k\to\infty,\ k\in K} \Delta f_l(x_k; d_k) = 0$ holds.

Proof. (i) The KKT conditions for problem LP(x_k) can be represented by

$$\begin{aligned}
&\nabla f(x_k) - A(x_k)^t y_{k+1} - z_{k+1} - \lambda_{k+1} + \mu_{k+1} = 0, \\
&A(x_k) d_k = 0, \quad \|d_k\|_\infty \leq 1 \\
&x_k + d_k \geq 0, \quad z_{k+1} \geq 0, \quad z_{k+1}^t (x_k + d_k) = 0, \\
&e + d_k \geq 0, \quad \lambda_{k+1} \geq 0, \quad \lambda_{k+1}^t (e + d_k) = 0, \\
&e - d_k \geq 0, \quad \mu_{k+1} \geq 0, \quad \mu_{k+1}^t (e - d_k) = 0,
\end{aligned} \tag{5}$$

where $z_{k+1}, \lambda_{k+1}, \mu_{k+1}$ are multipliers corresponding to the inequality constraints. By premultiplying the vector d_k^t, we have

$$d_k^t (\nabla f(x_k) - A(x_k)^t y_{k+1} - z_{k+1} - \lambda_{k+1} + \mu_{k+1}) = 0,$$

and we have

$$d_k^t \nabla f(x_k) + z_{k+1}^t x_k + \lambda_{k+1}^t e + \mu_{k+1}^t e = 0. \tag{6}$$

Since $d_k^t \nabla f(x_k) = \Delta f_l(x_k; d_k) = 0$, $z_{k+1}^t x_k \geq 0$, $\lambda_{k+1}^t e \geq 0$ and $\mu_{k+1}^t e \geq 0$, equation (6) yields

$$z_{k+1}^t x_k = 0, \quad \lambda_{k+1} = 0, \quad \mu_{k+1} = 0.$$

Therefore we obtain the result from (5).

(ii) In the same way as the proof of (i), we have equations (5) and (6). Since for $k \in K$, $d_k^t \nabla f(x_k) \to 0$, $z_{k+1}^t x_k \geq 0$, $\lambda_{k+1}^t e \geq 0$ and $\mu_{k+1}^t e \geq 0$, equation (6) yields

$$z_{k+1}^t x_k \to 0, \quad \lambda_{k+1} \to 0, \quad \mu_{k+1} \to 0.$$

Therefore we obtain the result.

(iii) Since $\Delta f_q(x_k; s_{T_k}) = 0$, the step $\tilde{s}_{T_k} = 0$ is also an optimal solution to subproblem $(QP_T(x_k))$. Thus there exist multipliers $\tilde{y}$, $\tilde{z}$, $\tilde{u}$ and $\tilde{v}$ corresponding to $\tilde{s}_{T_k}$ and they satisfy the following KKT conditions:

$$\begin{aligned}
&H_k \tilde{s}_{T_k} + \nabla f(x_k) - A(x_k)^t \tilde{y} - \tilde{z} - \tilde{u} + \tilde{v} = 0, \\
&A(x_k)\tilde{s}_{T_k} = 0, \quad x_k + \tilde{s}_{T_k} \geq 0, \quad -\Delta_{T_k} e \leq \tilde{s}_{T_k} \leq \Delta_{T_k} e, \\
&\tilde{z} \geq 0, \quad \tilde{u} \geq 0, \quad \tilde{v} \geq 0, \\
&\tilde{z}^t (x_k + \tilde{s}_{T_k}) = 0, \quad \tilde{u}^t (\Delta_{T_k} e + \tilde{s}_{T_k}) = 0, \quad \tilde{v}^t (\tilde{s}_{T_k} - \Delta_{T_k} e) = 0,
\end{aligned}$$

which yields

$$\nabla f(x_k) - A(x_k)^t \tilde{y} - \tilde{z} = 0, \quad x_k \geq 0, \quad \tilde{z} \geq 0, \quad \tilde{z}^t x_k = 0.$$

This implies that the point $(0, \tilde{y}, \tilde{z}, 0, 0)$ satisfies the KKT conditions for subproblem $(LP(x_k))$. Therefore the optimal value becomes $\Delta f_l(x_k; d_k) = \Delta f_l(x_k; 0) = 0$.

(iv) Let $\hat{s}_k = \Delta_{T_k} \dfrac{d_k}{\|d_k\|_\infty}$ and ξ_k be a positive number. Since $\xi_k \hat{s}_k$ is a feasible solution to $(QP_T(x_k))$ for ξ_k sufficiently small, we have

$$\begin{aligned}
\Delta f_q(x_k; s_{T_k}) &\leq \Delta f_q(x_k; \xi_k \hat{s}_k) \\
&= \xi_k \nabla f(x_k)^t \hat{s}_k + \frac{\xi_k^2}{2} (\hat{s}_k)^t H_k \hat{s}_k.
\end{aligned} \tag{7}$$

By choosing the sequence $\{\xi_k\}$ such that

$$\lim_{k \to \infty, k \in K} \frac{\Delta f_q(x_k; s_{T_k})}{\xi_k} = 0, \quad \lim_{k \to \infty, k \in K} \xi_k = 0,$$

equation (7) yields, for $k \in K$ sufficiently large,

$$\frac{\Delta f_q(x_k; s_{T_k})}{\xi_k} \leq \nabla f(x_k)^t \hat{s}_k + O(\xi_k) < 0.$$

Therefore, we have $\lim_{k \to \infty, k \in K} \Delta f_l(x_k; d_k) = 0$. □

The following lemma estimates $\Delta f_q(x_k; s_{T_k})$.

Lemma 2 *Assume that the point x_k satisfying $x_k \geq 0$ and $\|g(x_k)\| < \delta$ is given. Let Δ_{T_k} be sufficiently small. If $\Delta f_l(x_k; d_k) < 0$, then there exists a positive constant c_1 such that*

$$|\Delta f_q(x_k; s_{T_k})| \geq c_1 \|s_{T_k}\|_\infty.$$

Proof. By setting $\hat{v} = \Delta_{T_k} d_k$, the vector $\hat{v}$ becomes a feasible solution for subproblem $(QP_T(x_k))$, because $A(x_k)d_k = 0$ and $x_k + d_k \geq 0$ imply $A(x_k)\hat{v} = 0$ and $x_k + \hat{v} \geq 0$ respectively. Thus we have

$$\Delta f_q(x_k; s_{T_k}) \leq \Delta f_q(x_k; \hat{v}) \leq 0. \tag{8}$$

Since (8) holds for Δ_{T_k} sufficiently small, it follows that

$$\begin{aligned}
|\Delta f_q(x_k; s_{T_k})| &\geq |\Delta f_q(x_k; \hat{v})| \\
&= \left| \frac{1}{2} \hat{v}^t H_k \hat{v} + \nabla f(x_k)^t \hat{v} \right| \\
&\geq |\nabla f(x_k)^t \hat{v}| - \left| \frac{1}{2} \hat{v}^t H_k \hat{v} \right| \\
&\geq \frac{1}{2} |\nabla f(x_k)^t \hat{v}| \\
&\geq \frac{1}{2} |\nabla f(x_k)^t d_k| \|s_{T_k}\|_\infty.
\end{aligned}$$

Therefore the proof is complete. □

The following lemma guarantees that the procedures in Steps 4 and 5 are well defined.

Lemma 3 *Suppose that Assumption G holds. Then Step 4 always finds ρ_k which satisfies*

$$\Delta f_q(x_k; s_{\rho_k}) \leq \frac{1}{2} \Delta f_q(x_k; s_{T_k}).$$

Furthermore,

$$\|g(x_k + s_{\rho_k})\| < \delta$$

is satisfied for Δ_{T_k} sufficiently small.

Proof. We can assume $\Delta f_q(x_k; s_{T_k}) < 0$, because if $\Delta f_q(x_k; s_{T_k}) = 0$, then the algorithm terminates by (i) and (iii) of Lemma 1. Since the gradient vector, the matrix H_k and the radius Δ_{T_k} are uniformly bounded, we have

$$\begin{aligned}
&\Delta f_q(x_k; (1-\rho) s_{T_k} + \rho \bar{s}_k) - \frac{1}{2} \Delta f_q(x_k; s_{T_k}) \\
&= \frac{1}{2} \Delta f_q(x_k; s_{T_k}) + \rho(-\nabla f(x_k)^t s_{T_k} + \nabla f(x_k)^t \bar{s}_k - s_{T_k}^t H_k s_{T_k} + s_{T_k}^t H_k \bar{s}_k) \\
&\quad + \rho^2 \left(\frac{1}{2} s_{T_k}^t H_k s_{T_k} + \frac{1}{2} \bar{s}_k^t H_k \bar{s}_k - s_{T_k}^t H_k \bar{s}_k \right) \\
&\leq \frac{1}{2} \Delta f_q(x_k; s_{T_k}) + 2\rho(\sqrt{n} \|\nabla f(x_k)\| \|s_{T_k}\|_\infty + n\|H_k\| \|s_{T_k}\|_\infty^2) + 2n\rho^2 \|H_k\| \|s_{T_k}\|_\infty^2 \\
&\leq \frac{1}{2} \left(a \|s_{T_k}\|_\infty^2 \rho^2 + b \|s_{T_k}\|_\infty (1 + \|s_{T_k}\|_\infty) \rho + \Delta f_q(x_k; s_{T_k}) \right),
\end{aligned} \tag{9}$$

where a and b are positive constants independent of k. We note that

$$\begin{aligned}
&\frac{-b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty)+\sqrt{b^2\|s_{T_k}\|_\infty^2(1+\|s_{T_k}\|_\infty)^2-4a\|s_{T_k}\|_\infty^2\Delta f_q(x_k;s_{T_k})}}{2a\|s_{T_k}\|_\infty^2} \\
&= \frac{\sqrt{b^2\|s_{T_k}\|_\infty^2(1+\|s_{T_k}\|_\infty)^2-4a\|s_{T_k}\|_\infty^2\Delta f_q(x_k;s_{T_k})}-b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty)}{2a\|s_{T_k}\|_\infty^2} \cdot \\
&\quad \left(\frac{\sqrt{b^2\|s_{T_k}\|_\infty^2(1+\|s_{T_k}\|_\infty)^2-4a\|s_{T_k}\|_\infty^2\Delta f_q(x_k;s_{T_k})}+b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty)}{\sqrt{b^2\|s_{T_k}\|_\infty^2(1+\|s_{T_k}\|_\infty)^2-4a\|s_{T_k}\|_\infty^2\Delta f_q(x_k;s_{T_k})}+b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty)}\right) \\
&= \frac{2|\Delta f_q(x_k;s_{T_k})|}{\sqrt{b^2\|s_{T_k}\|_\infty^2(1+\|s_{T_k}\|_\infty)^2+4a\|s_{T_k}\|_\infty^2|\Delta f_q(x_k;s_{T_k})|}+b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty)} \\
&\geq \frac{|\Delta f_q(x_k;s_{T_k})|}{\gamma\|s_{T_k}\|_\infty}, \qquad (10)
\end{aligned}$$

where γ is a positive constant independent of k such that

$$\sqrt{b^2\|s_{T_k}\|_\infty^2(1+\|s_{T_k}\|_\infty)^2+4a\|s_{T_k}\|_\infty^2|\Delta f_q(x_k;s_{T_k})|}+b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty) \leq 2\gamma\|s_{T_k}\|_\infty.$$

Let $\bar{\rho}_k = \min\left(\frac{|\Delta f_q(x_k;s_{T_k})|}{\gamma\|s_{T_k}\|_\infty},\ 1\right)$. If we choose ρ such that $0 < \beta\bar{\rho}_k \leq \rho \leq \bar{\rho}_k$, then we have

$$a\|s_{T_k}\|_\infty^2\rho^2 + b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty)\rho + \Delta f_q(x_k;s_{T_k}) \leq 0.$$

Thus the procedure in Step 4 finds ρ_k which satisfies

$$\Delta f_q(x_k;s_{\rho_k}) \leq \frac{1}{2}\Delta f_q(x_k;s_{T_k}).$$

Since $A(x_k)s_{T_k} = 0$ and $A(x_k)s_k = -g(x_k)$, we have

$$\begin{aligned}
g(x_k) + A(x_k)s_{\rho_k} &= g(x_k) + A(x_k)\left((1-\rho_k)s_{T_k} + \rho_k \min\left\{\frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\} s_k\right) \\
&= \left(1 - \rho_k \min\left\{\frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\}\right) g(x_k).
\end{aligned}$$

We note that

$$\|s_{\rho_k}\|_\infty \leq (1-\rho_k)\|s_{T_k}\|_\infty + \rho_k \min\left\{\frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\}\|s_k\|_\infty = O(\|s_{T_k}\|_\infty)$$

and the bounds $\|s_k\|_\infty \leq \Delta_k \leq \bar{\Delta}$ and $\|s_{T_k}\|_\infty \leq \Delta_{T_k} \leq \bar{\Delta}$ yield

$$\min\left\{\frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\} \geq \min\left\{\frac{\|s_{T_k}\|_\infty}{\bar{\Delta}},\ 1\right\} = \frac{\|s_{T_k}\|_\infty}{\bar{\Delta}}.$$

Furthermore, since the algorithm does not terminate, we have $\Delta f_l(x_k; d_k) < 0$. Then $|\Delta f_q(x_k; s_{T_k})| \geq c_1 \|s_{T_k}\|_\infty$ holds for some positive constant c_1 by Lemma 2. Thus we have

$$\bar{\rho}_k = \min\left\{\frac{|\Delta f_q(x_k; s_{T_k})|}{\gamma \|s_{T_k}\|_\infty}, 1\right\} \geq \min\left\{\frac{c_1}{\gamma}, 1\right\}.$$

Using these facts, we obtain

$$\begin{aligned}
\|g(x_k + s_{\rho_k})\| &= \|g(x_k) + A(x_k)s_{\rho_k} + O(\|s_{\rho_k}\|^2)\| \\
&\leq \left(1 - \rho_k \min\left\{\frac{\|s_{T_k}\|_\infty}{\|s_k\|_\infty},\ 1\right\}\right) \|g(x_k)\| + O(\|s_{T_k}\|_\infty^2) \\
&< \left(1 - \rho_k \frac{\|s_{T_k}\|_\infty}{\bar{\Delta}}\right) \delta + O(\|s_{T_k}\|_\infty^2) \qquad (\text{since } \ \|g(x_k)\| < \delta) \\
&= \delta - \frac{\rho_k \delta}{\bar{\Delta}} \|s_{T_k}\|_\infty + O(\|s_{T_k}\|_\infty^2) \\
&\leq \delta - \beta \bar{\rho}_k \frac{\delta}{\bar{\Delta}} \|s_{T_k}\|_\infty + O(\|s_{T_k}\|_\infty^2) \qquad (\text{since } \ \beta \bar{\rho}_k \leq \rho_k) \\
&\leq \delta - \beta \frac{\delta}{\bar{\Delta}} \min\left\{\frac{c_1}{\gamma}, 1\right\} \|s_{T_k}\|_\infty + O(\|s_{T_k}\|_\infty^2). \qquad (11)
\end{aligned}$$

Therefore by (11), we obtain $\|g(x_k + s_{\rho_k})\| < \delta$ for Δ_{T_k} sufficiently small.

Therefore the proof is complete. □

The following lemma will be used in proving Theorem 2.

Lemma 4 *Suppose that Assumption G holds. If* $\liminf_{k \to \infty} |\Delta f_l(x_k; d_k)| = c_2 > 0$*, then*

$$\liminf_{k \to \infty} \rho_k > 0,$$

and there exists a positive number $\bar{\Delta}_T$ *independent of* k *such that the step* s_{ρ_k} *satisfies*

$$\|g(x_k + s_{\rho_k})\| < \delta$$

for any $\Delta_{T_k} \in (0, \bar{\Delta}_T)$*.*

Proof. We first note that if Δ_{T_k} is sufficiently small, the vector $\Delta_{T_k} d_k$ is a feasible solution to subproblem $(QP_T(x_k))$. Since $\liminf_{k \to \infty} |\Delta f_l(x_k; d_k)| > 0$ holds and the sequence $\{x_k\}$ remains in a compact set, we have

$$\begin{aligned}
\Delta f_q(x_k; s_{T_k}) &\leq \Delta f_q(x_k; \Delta_{T_k} d_k) \\
&= \frac{1}{2} \Delta_{T_k}^2 d_k^t H_k d_k + \Delta_{T_k} \Delta f_l(x_k; d_k) \\
&\leq \frac{1}{2} \Delta_{T_k} \Delta f_l(x_k; d_k) \\
&\leq \frac{1}{4} \Delta_{T_k}(-c_2) \\
&< 0
\end{aligned}$$

for Δ_{T_k} sufficiently small. In the same way as the proof of equation (10), the above inequalities yield

$$\frac{2|\Delta f_q(x_k; s_{T_k})|}{\sqrt{b^2\|s_{T_k}\|_\infty^2(1+\|s_{T_k}\|_\infty)^2 + 4a\|s_{T_k}\|_\infty^2|\Delta f_q(x_k; s_{T_k})|} + b\|s_{T_k}\|_\infty(1+\|s_{T_k}\|_\infty)^2}$$
$$\geq \frac{|\Delta f_q(x_k; s_{T_k})|}{\gamma\|s_{T_k}\|_\infty} \geq \frac{\Delta_{T_k} c_2}{4\gamma\|s_{T_k}\|_\infty} \geq \frac{c_2}{4\gamma}.$$

Let $\bar\rho = \min\left(\frac{c_2}{4\gamma},\ 1\right)$. Note that $\bar\rho$ is independent of the iteration count k. In Step 4, we can find ρ_k such that $0 < \beta\bar\rho \leq \rho_k \leq \bar\rho$. In the same way as the proof of equation (11), we have for some positive constant ξ

$$\begin{aligned} \|g(x_k + s_{\rho_k})\| &< \delta + \|s_{T_k}\|_\infty\left(-\frac{\rho_k\delta}{\bar\Delta} + O(\|s_{T_k}\|_\infty)\right) \\ &\leq \delta + \|s_{T_k}\|_\infty\left(-\beta\bar\rho\frac{\delta}{\bar\Delta} + \xi\Delta_{T_k}\right). \qquad (\text{since } \ \beta\bar\rho \leq \rho_k) \end{aligned}$$

Let $\bar\Delta_T = \dfrac{\beta\bar\rho\delta}{\xi\bar\Delta}$. We should note that $\bar\Delta_T$ is independent of the iteration count k. If we choose the radius Δ_{T_k} such that $0 < \Delta_{T_k} < \bar\Delta_T$, then we have

$$-\beta\bar\rho\frac{\delta}{\bar\Delta} + \xi\Delta_{T_k} < 0.$$

Thus the step s_{ρ_k} satisfies $\|g(x_k + s_{\rho_k})\| < \delta$ for any $\Delta_{T_k} \in (0, \bar\Delta_T)$.

Therefore the proof is complete. □

Now we present the global convergence theorem.

Theorem 2 *Suppose that Assumption G holds. Then Minimization Algorithm terminates in a finite number of iterations or there exists an accumulation point* $w_\infty = (x_\infty, y_\infty, z_\infty)^t$ *that satisfies*

$$\|g(x_\infty)\| \leq \delta, \quad \nabla_x L(w_\infty) = 0, \quad X_\infty Z_\infty e = 0, \quad x_\infty \geq 0, \quad z_\infty \geq 0, \tag{12}$$

for a given $\delta > 0$.

Proof. Suppose the infinite sequence is generated. Without loss of generality, we can assume that $\Delta f_q(x_k; s_{T_k}) < 0$ holds for all $k \geq 0$ from Lemma 1-(iii), which guarantees $\Delta f_q(x_k; s_{\rho_k}) < 0$ from Step 4 of the algorithm. Let $K_0 = \{k| \ \|g(x_k + s_{\rho_k})\| \geq \delta\}$. We define subsequences $K_1 \subset \{0, 1, \cdots\}$ and $K_2 \subset \{0, 1, \cdots\}$ that satisfy $K_1 \cup K_2 = \{0, 1, 2, \cdots\} \setminus K_0$ and $K_1 \cap K_2 = \emptyset$ by

$$\Delta f(x_k; s_{\rho_k}) > \frac{1}{4}\Delta f_q(x_k; s_{\rho_k}), \ \ k \in K_1, \tag{13}$$

$$\Delta f(x_k; s_{\rho_k}) \leq \frac{1}{4}\Delta f_q(x_k; s_{\rho_k}),\ \ k \in K_2. \tag{14}$$

(i) Suppose that K_1 is an infinite sequence.

(i-a) Suppose that $\liminf_{k\to\infty, k\in K_1} \Delta_{T_k} = 0$. Then there exists an infinite set $K_1' \subset K_1$ such that $\Delta_{T_k} \to 0$ for $k \in K_1'$. Note that $\|s_{\rho_k}\| \to 0$ for $k \in K_1'$. Since equation (13) yields

$$\frac{1}{4}\Delta f_q(x_k; s_{\rho_k}) < \Delta f(x_k; s_{\rho_k}) = \Delta f_q(x_k; s_{\rho_k}) + \mathrm{O}\left(\|s_{\rho_k}\|_\infty^2\right),$$

there exists a positive constant c_3 such that

$$-\Delta f_q(x_k; s_{\rho_k}) \leq c_3\|s_{\rho_k}\|_\infty^2 \leq c_3\Delta_{T_k}^2. \tag{15}$$

Assume that $\liminf_{k\to\infty, k\in K_1'} \Delta f_l(x_k; d_k) < 0$. Without loss of generality, we can assume that $\lim_{k\to\infty, k\in K_1'} \Delta f_l(x_k; d_k) < 0$. Then $\hat{s}_k = \Delta_{T_k}\dfrac{d_k}{\|d_k\|_\infty}$ is a feasible solution to $(QP_T(x_k))$. Since the matrix $\{\nabla^2 f(x_k)\}$ is uniformly bounded, we have

$$\begin{aligned} 2\Delta f_q(x_k; s_{\rho_k}) &\leq \Delta f_q(x_k; s_{T_k}) \\ &\leq \Delta f_q(x_k; \hat{s}_k) \\ &\leq \Delta_{T_k}\left(\frac{\Delta f_l(x_k; d_k)}{\|d_k\|_\infty}\right) + c_4\Delta_{T_k}^2 \end{aligned}$$

for a positive constant c_4. Thus there exists a positive constant c_5 such that

$$-2\Delta f_q(x_k; s_{\rho_k}) \geq \Delta_{T_k}\left(\frac{-\Delta f_l(x_k; d_k)}{\|d_k\|_\infty}\right) - c_4\Delta_{T_k}^2 \geq c_5\Delta_{T_k}, \tag{16}$$

for $k \in K_1'$ sufficiently large. This contradicts equation (15). This implies that

$$\lim_{k\to\infty, k\in K_1'} \Delta f_l(x_k; d_k) = 0.$$

Thus it follows from Lemma 1-(ii) that any accumulation point of the sequence $\{w_k\}$ satisfies equations (12).

(i-b) Suppose that $\liminf_{k\to\infty, k\in K_1} \Delta_{T_k} > 0$. Then the condition $\Delta f(x_k; s_{\rho_k}) \leq \frac{3}{4}\Delta f_q(x_k; s_{\rho_k})$ must be satisfied infinitely many times for $k \notin K_1$ and this case corresponds to (ii) below.

(ii) Suppose that K_2 is an infinite sequence.

(ii-a) Suppose that there exists an infinite sequence $K_2' \subset K_2$ such that $\liminf_{k\to\infty, k\in K_2'} \Delta_{T_k} > 0$. Since $\{f(x_k)\}$ is bounded below and decreasing, and $\Delta f(x_k; s_{\rho_k}) \leq 0$ for $k \in K_2$, we have

$$f(x_{k+1}) - f(x_k) = \Delta f(x_k; s_{\rho_k}) \to 0, \quad k \in K_2$$

and thus $\Delta f_q(x_k; s_{\rho_k}) \to 0,\ k \in K_2$, from (14), which implies that $\Delta f_q(x_k; s_{T_k}) \to 0,\ k \in K_2$ by (4). It follows from Lemma 1-(iv) that $\lim_{k\to\infty, k\in K_2'} \Delta f_l(x_k; d_k) = 0$. Thus it follows from Lemma 1-(ii) that any accumulation point of the sequence $\{w_k\}$ satisfies equations (12).

(ii-b) Suppose that $\lim_{k\to\infty, k\in K_2} \Delta_{T_k} = 0$ holds. Assume that any accumulation point does not satisfy conditions (12), which implies $\liminf_{k\to\infty, k\in K_2} |\Delta f_l(x_k; d_k)| = c_6 > 0$ by Lemma 1 - (ii). It follows from Lemma 4 that the numbers of each consecutive iteration that belongs to K_0 are uniformly bounded. Thus K_1 is an infinite sequence and this case reduces to Case (i). For Case (i-a), we obtain the desired result. On the other hand, Case (i-b) does not hold, because if this is the case, then we have

$$\liminf_{k\to\infty, k\in K_1} \Delta_{T_k} > 0 \quad \text{and} \quad \lim_{k\to\infty, k\in K_2} \Delta_{T_k} = 0,$$

which is a contradiction. Hence there exists an accumulation point that satisfies conditions (12).

Therefore, the proof of the theorem is complete. □

4.3 Global convergence of Main Algorithm

To end this section, we present the global convergence property of Main Algorithm as the following theorem.

Theorem 3 *Suppose that Assumption G holds. Then Main Algorithm terminates in a finite number of iterations or any accumulation point of a generated sequence $\{w_k\}$ satisfies the KKT conditions.*

Proof. The conditions $\|r(w_k)\|_* \le \delta_k$, $x_k \ge 0$ and $z_k \ge 0$ in Step 2 are satisfied at each iteration and the sequence $\{\delta_k\}$ approaches zero, the result immediately follows. □

5 Superlinear convergence of Main Algorithm

In this section, we analyze the rate of convergence of Main Algorithm under the assumption that the sequence $\{w_k\}$ converges to a KKT point w^*. We will show that in Main Algorithm, Restoration Algorithm terminates in one iteration with a unit step size and Minimization Algorithm is skipped. The local convergence property of the standard SQP method enables us to obtain the superlinear convergence property of Main Algorithm. Let x_k be the kth iterate of Main Algorithm. Thus x_k is used as an initial point of Restoration Algorithm. In what follows, we add a superscript "r" to scalars and vectors dealt in Restoration Algorithm in order to distinguish from those used in Main Algorithm. Furthermore, we omit the iteration count in Restoration Algorithm, because the first step is only dealt with in Restoration Algorithm. Suppose that w_k is sufficiently close to a KKT point w^*. We should note that the initial trust-region radius Δ^r of Restoration Algorithm that starts from w_k is uniformly bounded away from zero. Therefore, the trust-region constraint $\|s^r\|_\infty \leq \Delta^r$ becomes inactive, because the magnitude of $\|s^r\|_\infty$ is sufficiently small and the trust-region radius Δ^r is bounded away from zero.

We first give the following lemma.

Lemma 5 *Suppose that the sequence $\{w_k\}$ generated by Main Algorithm converges to a KKT point w^*. Assume that the linear independence constraint qualification, the strict complementarity condition and the second-order sufficient condition are satisfied at w^*. Let s^r, y^r_+ and z^r_+ be an optimal solution and corresponding multipliers for $QP_R(x_k)$, respectively. Suppose that the matrices G_k are uniformly positive definite on the null spaces of the matrices whose rows are $\nabla g_i(x_k)^t, i = 1, \ldots, m$ and the gradient vectors of active inequality constraints. Suppose that the matrices G_k are uniformly bounded and have uniformly bounded inverses, and that the matrices G_k satisfy*

$$\lim_{k\to\infty} \frac{\|P_k(G_k - \nabla_x^2 L(w^*))((x_k + s^r) - x_k)\|}{\|(x_k + s^r) - x_k\|} = 0,$$

where P_k is the orthogonal projection matrix onto the tangent space to the equality and active inequality constraints at x_k. Then the following holds for k sufficiently large:

$$\|r(x_k + s^r, y^r_+, z^r_+)\| \leq \beta_k \|r(w_k)\|,$$

where $\{\beta_k\}$ is a positive sequence such that $\lim_{k\to\infty} \beta_k = 0$. □

The proof can be found in [1]. We note that $x_k + s^r \geq 0$, $z^r_+ \geq 0$ and $(z^r_+)_i (x_k + s^r)_i = 0, i = 1, \ldots, n$ hold. By using the preceding lemma, we have the following theorem.

Theorem 4 *Suppose that all the assumptions of Lemma 5 hold. Then for k sufficiently large, Restoration Algorithm terminates with $w_{k+\frac{1}{2}} = (x_k + s^r, y^r_+, z^r_+)$, and Main Algorithm adopts $w_{k+1} = w_{k+\frac{1}{2}}$ in Step 2 and skips Step 3. Therefore, the sequence $\{w_k\}$ generated by Main Algorithm satisfies the relation*

$$\|r(w_{k+1})\| \leq \beta_k \|r(w_k)\|. \tag{17}$$

Proof. Since

$$\|r(w)\|_* = \max(\|\nabla_x L(w)\|, \|g(x)\|, \|XZe\|) \leq \|r(w)\|$$

and

$$\|r(w)\|^2 = \|\nabla_x L(w)\|^2 + \|g(x)\|^2 + \|XZe\|^2 \leq 3\|r(w)\|_*^2,$$

we have

$$\frac{1}{\sqrt{3}}\|r(w)\| \leq \|r(w)\|_* \leq \|r(w)\|. \tag{18}$$

Equation (18) and Lemma 5 yield

$$\begin{aligned}
\|g(x_k + s^r)\| &\leq \|r(x_k + s^r, y_+^r, z_+^r)\|_* \\
&\leq \|r(x_k + s^r, y_+^r, z_+^r)\| \\
&\leq \beta_k \|r(w_k)\| \\
&\leq \sqrt{3}\beta_k \|r(w_k)\|_* \\
&= \left(\frac{\sqrt{3}\beta_k}{\tau}\right)\delta_k \\
&< \delta_k,
\end{aligned}$$

for k sufficiently large. Since the relations

$$\|g(x_k + s^r)\| < \delta_k, \quad x_k + s^r \geq 0 \quad \text{and} \quad z_+^r \geq 0$$

hold, we have

$$w_{k+\frac{1}{2}} = (x_k + s^r, y_+^r, z_+^r).$$

Furthermore the fact $\|r(x_k + s^r, y_+^r, z_+^r)\|_* < \delta_k$ yields

$$w_{k+1} = w_{k+\frac{1}{2}},$$

which implies that Step 3 is skipped. Therefore, by Lemma 5, the proof is complete. □

Equation (17) implies that our method converges superlinearly to a KKT point.

6 Concluding remarks

In this paper, we have proposed a new trust-region SQP method. Our method consists of Restoration Algorithm and Minimization Algorithm, and the method does not use a penalty function. We have proved the global and superlinear convergence property of the present method. We note that our method is different from the filter method proposed by Fletcher et al. We emphasize that our method avoids the Maratos effect.

In Step 1 of Main Algorithm, we have specified the way of choice of the parameter δ_k. This choice of δ_k plays an important role in showing the superlinear convergence property of our method, but the condition that $\{\delta_k\}$ approaches zero is sufficient to prove the global convergence property.

References

[1] P. T. Boggs and J. W. Tolle, Sequential quadratic programming, in *Acta Numerica 1995*, Cambridge University Press, pp. 1-51, 1995.

[2] R. Fletcher, N. I. M. Gould, S. Leyffer, Ph. L. Toint and A. Wächter, Global convergence of a trust-region SQP-filter algorithm for general nonlinear programming, *SIAM Journal on Optimization*, 13 (2002), pp. 635-659.

[3] R. Fletcher and S. Leyffer, Nonlinear programming without a penalty function, *Mathematical Programming*, 91 (2002), pp. 239-269.

[4] M. Ulbrich and S. Ulbrich, Non-monotone trust-region methods for nonlinear equality constrained optimization without a penalty function, *Mathematical Programming*, 95 (2003), pp. 103-135.

[5] H. Yamashita, *A globally convergent quasi-Newton method for equality constrained optimization that does not use a penalty function*, Technical Report, Mathematical Systems Inc., Tokyo, Japan, June 1979 (revised September 1982).