Trust Region Method of Conic Model for Linearly Constrained Optimization

by

Wenyu Sun, Jinyun Yuan and Ya-xiang Yuan

Report No.**ICM-96-003** January 1996

Trust Region Method of Conic Model for Linearly Constrained Optimization *

Wenyu Sun,[†] Jinyun Yuan[‡] and Ya-xiang Yuan[§]

June 4, 2007

Abstract

In this paper we present a trust region method of conic model for linearly constrained optimization problems. We discuss trust region approachs with conic model subproblems. Some equivalent variation properties and optimality conditions are given. A trust region algorithm based on conic model is constructed. Global convergence of the method is proved.

Key words. trust region method, conic model, constrained optimization.
AMS (MOS) subject classification. 65k, 90c

1. Introduction

Trust region methods have very nice global and local convergence properties, and it has been shown that they are very effective and robust for solving

*This work was supported by CNPq, Brazil and the National Natural Science Foundation of China.

[†]Department of Mathematics, Nanjing University, Nanjing 210093, China

[‡]Department of Mathematics, The Federal University of Parana, Curitiba 81531-990, PR, Brazil

[§]LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing 100080, China

unconstrained and constrained optimization problems (for exampel, see [2], [3], [4], [6] , [8], [9], [11], [12], [13], [15], [17] and [22]). Conic model methods, a generalization of quadratic model methods, possess more degree of freedom, can incorporate more information in the iterations, and provide both a powerful unifying theory and a effective means for optimization problems [1] [4] [5] [10] [16] [19] [21].

In [4], we present a trust region method of conic model for unconstrained optimization problems. As a continuing work, in this paper, we describe a trust region method of conic model to solve linearly constrained optimization problem

$$\min \quad f(x) \tag{1.1}$$
$$\text{s.t.} \quad A^T x = b, \tag{1.2}$$

where $f : R^n \to R$ is continuously differentiable, $A \in R^{n\times m}, x \in R^n, b \in R^m, \text{rank}(A) = m$. Our method is iterative, and the trust region subproblem solved in each iteration is the minimization of a conic model subject to the linear constraints and an additional trust region constraint.

Normally , numerical methods for solving optimization problem (1.1)-(1.2) are reduced gradient method, projected gradient method and reduced quasi-Newton method which are based on quadratic model. Using null space techniques, the constrained problem (1.1)-(1.2) can be transformed to an unconstrained problem. In order to incorporate more useful interpolation information in constructing subproblems, Davidon [5] suggested a new model – conic model. A typical conic model for unconstrained optimization is as follows:

$$\psi(s) = f_k + \frac{g_k^T s}{1 - a^T s} + \frac{1}{2}\frac{s^T A_k s}{(1 - a^T s)^2}, \tag{1.3}$$

where $f_k = f(x_k), g_k = \nabla f(x_k), A_k \in R^{n\times n}$ is a symmetric matrix, the vector $a \in R^n$ is a vector satisfying $1 - a^T s > 0$. If $a = 0, \psi(s)$ is quadratic.

The conic model (1.3) can be also written as the following form of the collinear conic model:

$$\psi(s) = f_k + g_k^T w + \frac{1}{2} w^T A_k w \tag{1.4}$$
$$s = \frac{w}{1 + a^T w}. \tag{1.5}$$

It follows from (1.4)-(1.5) that

$$s = \frac{-A_k^{-1} g_k}{1 - a^T A_k^{-1} g_k}$$

is a minimizer of $\psi(s)$ if A_k is positive definite.

Sorensen [16] discussed collinear scaling methods for unconstrained optimization. For the scaling function

$$\phi_{k+1}(w) = f(\bar{x}(w)) = f(x_{k+1} + \frac{w}{1 + h_{k+1}^T w}),$$

the corresponding quadratic model is

$$\psi_{k+1}(w) = \phi_{k+1}(0) + \phi'_{k+1}(0)w + \frac{1}{2} w^T B_{k+1} w,$$

which satisfies the following interpolation conditions

$$\psi_{k+1}(0) = \phi_{k+1}(0),\ \psi'_{k+1}(0) = \phi'_{k+1}(0),$$

$$\psi_{k+1}(-v) = \phi_{k+1}(-v),\ \psi'_{k+1}(-v) = \phi'_{k+1}(-v),$$

where $v \in R^n$ is chosen such that $1 - h_{k+1}^T v > 0$.

Di and Sun [4] consider a trust region method of conic model for unconstrained optimization. They give the following model

$$\min \quad \psi(s) = f_k + \frac{g_k^T s}{1 - a^T s} + \frac{1}{2} \frac{s^T B_k s}{(1 - a^T s)^2} \tag{1.6}$$

$$\text{s.t.} \quad \|Ds\| \leq \Delta_k \tag{1.7}$$

or equivalently

$$\min \quad f_k + g_k^T J_k w + \frac{1}{2} w^T B_k w \tag{1.8}$$

$$\text{s.t.} \quad s = \frac{J_k w}{1 + h^T w},\ \|Ds\| \leq \Delta_k. \tag{1.9}$$

They construct a trust region algorithm based on the above model, and give convergence analyses.

In this paper we generalize the trust region method of conic model for unconstrained optimization to solve linearly constrained optimization problem (1.1)-(1.2). In Section 2, the motivation of a detailed description of our method are given. Convergence analyses of the new algorithm are presented in Section 3.

2. Motivation and Description of the Algorithm

Assume that the current point x_k is feasible, namely $A^T x_k = b$, it is easy to see that the constrained condition is equivalent to $A^T s = 0$ if we let $x = x_k + s$. Therefore it is reasonable to use the following subproblem:

$$\min \quad \psi_k(s) \tag{2.1}$$

$$\text{s.t.} \quad A^T s = 0 \tag{2.2}$$

$$\|s\| \leq \Delta_k \tag{2.3}$$

where

$$\psi_k(s) = \frac{g_k^T s}{1 - h^T s} + \frac{1}{2} \frac{s^T B_k s}{(1 - h^T s)^2}, \tag{2.4}$$

and $g_k = g(x_k) = \nabla f(x_k)$, B_k is an approximation of the Hessian matrix $\nabla^2 f(x_k)$ and $h \in R^n$ is a horizon vector such that $1 - h^T s > 0$.

Comparing the above subproblem with (1.8)-(1.9), one can easily see that we have choose $D = I$ and $J_k = I$ for all k. It should be pointed out our results in the paper can be extented to general D and J_k, and we make these special choices for the convenience of convergence analyses. Though theoretical analyses are nearly identical for general D and J_k, numerical performances of the algorithms will vary for different choices of D and J_k.

We will use a null space technique to handle the constraint (2.2). Let $Y \in R^{n \times m}$ and $Z \in R^{n \times (n-m)}$ be two matrices that satisfy

$$A^T Y = I, \; A^T Z = 0, \; Z^T Z = I$$

with rank$(Z) = n - m$. For example, Y and Z can be obtained from the QR decomposition of A. Assume that

$$A = Q \begin{bmatrix} R \\ 0 \end{bmatrix} = [Q_1 \; Q_2] \begin{bmatrix} R \\ 0 \end{bmatrix}$$

where Q is $n \times n$ orthogonal matrix, $R \in R^{m\times m}$ is a nonsingular upper triangular matrix, and Q_1 and Q_2 are $n \times m$ and $n \times (n-m)$ matrices respectively. We can choose

$$Y = (A^+)^T = Q_1 R^{-T},$$

$$Z = Q_2,$$

where A^+ is a Moore-Penrose inverse (see [6] [7]). Since A is a column full-rank matrix, then $A^+ = (A^T A)^{-1} A^T$. Obviously, the columns of Z form an orthogonal basis for the null space of A^T. Therefore condition (2.2) reduced to $s = Zu$, where $u \in R^{n-m}$. Therefore our subproblem (2.1)-(2.3) can be rewritten as

$$\begin{aligned} \min \quad & \hat{\psi}_k(u) = \frac{\hat{g}_k^T u}{1 - \hat{h}^T u} + \frac{1}{2}\frac{u^T \hat{B}_k u}{(1-\hat{h}^T u)^2} && (2.5)\\ \text{s.t.} \quad & \|u\| \le \Delta_k, && (2.6) \end{aligned}$$

where $\hat{g}_k = Z^T g_k$ is a reduced gradient, $\hat{B}_k = Z^T B_k Z$ is a reduced Hessian approximation and $\hat{h} = Z^T h$ is a reduced horizon vector. In fact the above subproblem (2.5)-(2.6) is a conic trust region subproblem for the unconstrained optimization

$$\min_{\hat{x}\in \Re^{n-m}} \hat{f}(\hat{x}) \tag{2.7}$$

where $\hat{f}(\hat{x}) = f(x_k + Z\hat{x})$. Problem (2.5)-(2.6) can be solved by techniques given by Di and Sun [4]. It is easy to see that a solution u_k of (2.5)-(2.6) satisfies

$$\begin{aligned} (\hat{B}_k - \hat{g}_k \hat{h}^T + \mu_k I) u_k \quad &= \quad -\hat{g}_k + \mu_k \Delta_k^2 \hat{h}, && (2.8)\\ \mu_k(\|u_k\| - \Delta_k) = 0, & && (2.9) \end{aligned}$$

which, in fact, is the first order optimality condition for problem (2.5)-(2.6), where $\mu_k \ge 0$ is a Kuhn-Tucker multiplier.

Lemma 2.1 *Conic model subproblem (2.1) with trust region in constrained form can be transformed to a quadratic model subproblem with trust region in unconstrained form.*

Proof. Let $w = \frac{s}{1-h^Ts}$ (i.e., $s = \frac{w}{1+h^Tw}$), then (2.1) becomes

$$\min \quad g_k^T w + \frac{1}{2} w^T B_k w \tag{2.10}$$
$$\text{s.t.} \quad A^T w = 0 \tag{2.11}$$
$$\| \frac{w}{1+h^T w} \| \leq \Delta_k \tag{2.12}$$

which is to minimize a quadratic function subject to linear constraints and a conic type trust region constraint. This trust region always lies in $\{\bar{x} \in R^n | \ 1 + h^T(\bar{x} - \bar{x}_k) > 0\}$ for any Δ_k and $\bar{x} = \bar{x}_k + w$.

Set $w = Z\hat{w}, \hat{g}_k = Z^T g_k, \hat{B}_k = Z^T B_k Z, \hat{h} = Z^T h$, then (2.10)-(2.12) becomes

$$\min \quad \hat{g}_k \hat{w} + \frac{1}{2} \hat{w}^T \hat{B}_k \hat{w} \tag{2.13}$$
$$\text{s.t.} \quad \frac{\hat{w}^T \hat{w}}{(1 + \hat{h}^T \hat{w})^2} \leq \Delta_k^2. \tag{2.14}$$

Note that the conic trust region (2.14) can be written as an ellipsoid trust region. In fact, (2.14) is equivalent to

$$\hat{w}^T \hat{w} \leq (1 + \hat{h}^T \hat{w})^2 \Delta_k^2. \tag{2.15}$$

Let Q be the orthogonal rotation matrix such that

$$Q\hat{h} = \|\hat{h}\| e_1, \tag{2.16}$$

where $e_1 = (1, 0, \ldots, 0)^T$. It can be shown that (2.15) is equivalent to

$$\theta(\bar{w}_1 - \omega)^2 + \bar{w}_2^2 + \cdots + \bar{w}_n^2 \leq \bar{\Delta}_k^2, \tag{2.17}$$

where $\{\bar{w}_i, i = 1, \cdots, n\}$ are components of the vector $\bar{w} = Q\hat{w}$, and

$$\theta = 1 - \|\hat{h}\|^2 \Delta_k^2, \ \omega = \frac{\|\hat{h}\| \Delta_k^2}{\theta}, \ \bar{\Delta}_k = \frac{\Delta_k}{\sqrt{\theta}}. \tag{2.18}$$

Define

$$\hat{z} = \bar{w} - \omega e_1, \ V = \text{diag}(\theta, 1, \cdots, 1), \tag{2.19}$$

(2.17) reduces to

$$\hat{z}^T V \hat{z} \le \bar{\Delta}_k^2. \tag{2.20}$$

Therefore, subproblem (2.13)-(2.14) becomes

$$\min \quad \bar{g}_k^T \hat{z} + \frac{1}{2}\hat{z}^T \bar{B}_k \hat{z} \tag{2.21}$$

$$\text{s.t.} \quad \hat{z}^T V \hat{z} \le \bar{\Delta}_k^2, \tag{2.22}$$

where $\bar{g}_k = Q\hat{g}_k, \bar{B}_k = Q\hat{B}_k Q^T$. Setting $z = V^{\frac{1}{2}}\hat{z}$ (2.21)Φ-(2.22) yield

$$\min \quad \tilde{g}_k^T z + \frac{1}{2} z^T \tilde{B}_k z \tag{2.23}$$

$$\text{s.t.} \quad \|z\| \le \bar{\Delta}_k, \tag{2.24}$$

where $\tilde{g}_k = V^{\frac{1}{2}}\bar{g}_k, \tilde{B}_k = V^{\frac{1}{2}}\bar{B}_k V^{\frac{1}{2}}$. (2.23)-(2.24) just is the desired quadratic model subproblem with trust region in unconstrained form, which can be solved by algorithms by [9]. □

From the above analyses, it can be seen that five supbproblems, (2.1)-(2.2), (2.5)-(2.6), (2.10)-(2.12), (2.13)-(2.14), and (2.23)-(2.24), are equivalent. Therefore in the algorithm we can solve any of them. Since our subproblem is based on conic model, these models possess more degree of freedom to incorporate interpolation information in iterative processes.

In the following we give a description of our algorithm. Reduced quasi-Newton methods are used to update the conic model. In the reduced form of updating $\hat{B}_k$, updating formula is written as

$$\hat{B}_{k+1} = U(\hat{B}_k, v_k, r_k),$$

where updating relation U is BFGS or DFP formula. The conic model satisfies the following generalized quasi-Newton equation:

$$\hat{B}_{k+1} v_k = r_k, \quad v_k = \gamma_k u_k, \quad \hat{h}_{k+1}^T v_k = 1 - \gamma_k,$$

where

$$r_k = \hat{g}_{k+1} - \frac{1}{\gamma_k}[I + \hat{h}_{k+1} u_k^T]\hat{g}_k,$$

$$\hat{h}_{k+1} = \frac{1-\gamma_k}{\gamma_k u_k^T p_k} p_k,$$

where $p_k \in R^n$ such that $u_k^T p_k \neq 0$.

Generally, we have two choices for p_k.

(1) Set $p_k = \hat{g}_k$, then

$$\hat{h}_{k+1} = \frac{1-\gamma_k}{\gamma_k u_k^T \hat{g}_k} \hat{g}_k \triangleq \alpha_k \hat{g}_k, \ r_k = y_k/\gamma_k,$$

where $y_k = \gamma_k \hat{g}_{k+1} - \frac{1}{\gamma_k}\hat{g}_k$.

(2) Set $p_k = \hat{g}_{k+1}$, then

$$\hat{h}_{k+1} = \frac{1-\gamma_k}{\gamma_k u_k^T \hat{g}_{k+1}} \hat{g}_{k+1} \triangleq \alpha_{k+1} \hat{g}_{k+1},$$

$$r_k = \beta_k \hat{g}_{k+1} - \frac{1}{\gamma_k}\hat{g}_k,$$

where

$$\beta_k = 1 - \frac{1-\gamma_k}{\gamma_k^2} \frac{u_k^T \hat{g}_k}{u_k^T \hat{g}_{k+1}}.$$

In the following algorithm the ratio of the actual reduction and the predicted reduction is defined as

$$\rho_k = \frac{Ared_k}{Pred_k} = \frac{f(x_k) - f(x_k + s_k)}{-\hat{\psi}_k(u_k)}.$$

Note that $s_k = 0$ if and only if x_k is a Kuhn-Tucker point of (1.1).

Algorithm 2.2 *(Conic Trust Region Algorithm for Linear Constrained Optimization)*

Step 0. Given a starting point x_0, an initial approximation to the reduced Hessian $\hat{B}_0 \in R^{n\times n}$, an initial trust region radius Δ_0 and $\epsilon > 0$. Given Z satisfying $A^T Z = 0$ with rank$(Z) = n - m$. Set $\mu \in [0,1), \eta \in (\mu, 1), 0 < \xi_0 < \xi_1 < 1 < \xi_2$. Set $k = 0$.

Step 1. Compute $f(x_k), g(x_k)$ and $\hat{g}_k = Z^T g_k$. If $\|\hat{g}_k\| \leq \epsilon$, stop.

Step 2. Solve the trust region subproblem (2.5) of conic model for u_k and s_k.

Step 3. Compute

$$\rho_k = \frac{f(x_k) - f(x_k + s_k)}{-\hat{\psi}_k(u_k)}.$$

Step 4. Set

$$x_{k+1} = \begin{cases} x_k + s_k & \text{if } \rho_k > \mu, \\ x_k & \text{otherwise ;} \end{cases} \tag{2.25}$$

and let the new trust region bound satisfy

$$\Delta_{k+1} \in [\Delta_k, \xi_2 \Delta_k], \text{ if } \rho_k \geq \eta \tag{2.26}$$

$$\Delta_{k+1} \in [\xi_0 \|s_k\|, \xi_1 \Delta_k], \text{ if } \rho_k < \eta. \tag{2.27}$$

Step 5. Update $\hat{B}_k$.

$$\begin{aligned}
&\varrho_k = [(\hat{f}(\hat{x}_k) - \hat{f}(\hat{x}_{k+1}))^2 - \hat{g}_k^T u_k \hat{g}_{k+1}^T u_k]^{\frac{1}{2}}, \\
&\gamma_k = -\tfrac{\hat{g}_k^T u_k}{\hat{f}(\hat{x}_k) - \hat{f}(\hat{x}_{k+1}) + \varrho_k}, \\
&v_k = \gamma_k u_k, \\
&y_k = \gamma_k \hat{g}_{k+1} - \tfrac{1}{\gamma_k} \hat{g}_k, \\
&r_k = y_k / \gamma_k, \\
&\alpha_k = \tfrac{1 - \gamma_k}{\gamma_k u_k^T \hat{g}_k}, \\
&\hat{B}_{k+1} = U(\hat{B}_k, v_k, r_k), \\
&\hat{h}_{k+1} = \alpha_k \hat{g}_k.
\end{aligned}$$

Step 6. $k := k + 1$, *go to Step 1.*

In our algorithm, we can allow $\mu = 0$. By setting $\mu = 0$, the algorithm has the nice property that any "better" point will be accepted. However the convergence results are not the same for the case $\mu = 0$ and $\mu > 0$.

3. Global Convergence

In this section, we give the convergence results of our algorithm given in the previous section.

The following lemma is important for convergence analyses of trust region algorithms, which is a generalization of a result by Powell [11] for unconstrained optimization.

Consider

$$\begin{aligned} \min \quad & \hat{\psi}_k(u) = \frac{\hat{g}_k^T u}{1 - \hat{h}^T u} + \frac{1}{2} \frac{u^T \hat{B}_k u}{(1 - \hat{h}^T u)^2} \\ \text{s.t.} \quad & \|u\| \leq \Delta_k. \end{aligned} \tag{3.1}$$

Lemma 3.1 *If u_k is the solution of (3.1) and if $\|\cdot\|$ is the l_2 norm, then*

$$Pred_k(u_k) = -\hat{\psi}_k(u_k) \geq \frac{1}{2}\|\hat{g}_k\| \min\{\tilde{\Delta}_k, \frac{\|\hat{g}_k\|}{\|\hat{B}_k\|}\}, \tag{3.2}$$

where

$$\tilde{\Delta}_k = \frac{\Delta_k}{1 + \Delta_k \hat{h}^T \hat{g}_k / \|\hat{g}_k\|}. \tag{3.3}$$

Proof. Let $v = \frac{u}{1 - \hat{h}^T u}$, then $u = \frac{v}{1 + \hat{h}^T v}$ and (3.1) becomes

$$\hat{g}_k^T v + \frac{1}{2} v^T \hat{B}_k v.$$

Let $v(\tau) = -\tau \frac{\hat{g}_k}{\|\hat{g}_k\|}, \tau > 0$, then

$$u(\tau) = \frac{-\tau \hat{g}_k}{\|\hat{g}_k\| - \tau \hat{h}^T \hat{g}_k} \tag{3.4}$$

and

$$\phi(\tau) = \phi(u(\tau)) = -\tau\|\hat{g}_k\| + \frac{1}{2}\tau^2 m_k, \tag{3.5}$$

where

$$m_k = \frac{\hat{g}_k^T \hat{B}_k \hat{g}_k}{\|\hat{g}_k\|^2}.$$

$\tau^* = \frac{\|\hat{g}_k\|}{m_k}$ is the minimizer of (3.5).

If $\|u(\tau^*)\| \leq \Delta_k$, then

$$\phi(\tau^*) = -\frac{1}{2}\frac{\|\hat{g}_k\|^2}{m_k} \leq -\frac{1}{2}\frac{\|\hat{g}_k\|^2}{\|\hat{B}_k\|}. \tag{3.6}$$

If $\|u(\tau^*)\| = \frac{\tau^* \|\hat{g}_k\|}{|\|\hat{g}_k\| - \tau^* \hat{h}^T \hat{g}_k|} \geq \Delta_k$, we choose τ_0 such that

$$\frac{\tau_0 \|\hat{g}_k\|}{|\|\hat{g}_k\| - \tau_0 \hat{h}^T \hat{g}_k|} = \Delta_k,$$

i.e.,

$$\tau_0 = \frac{\Delta_k}{|1 + \Delta_k \hat{h}^T \hat{g}_k / \|\hat{g}_k\||} \equiv \tilde{\Delta}_k,$$

then $\tau^* > \tau_0$, that is

$$\frac{\|\hat{g}_k\|}{m_k} \geq \tilde{\Delta}_k.$$

Then

$$\begin{aligned} \phi(\tau^*) &\leq -\frac{1}{2}\|\hat{g}_k\|\tilde{\Delta}_k(2 - \tilde{\Delta}_k \frac{m_k}{\|\hat{g}_k\|}) \qquad (3.7) \\ &\leq -\frac{1}{2}\|\hat{g}_k\|\tilde{\Delta}_k. \end{aligned}$$

Since $\hat{\psi}_k(u_k) \leq \phi(\tau^*)$, the result follows from (3.6) and (3.7). □

The condition (3.2) is quite general. First, it allows the step u_k to be obtained by several methods. Second, the reduced horizon vector $\hat{h}$ can be chosen as long as it satisfies $1 - \hat{h}^T u > 0$. In above algorithm we use $\hat{h} = \hat{g}_k$. Third, it allows choosing l_1, l_2 or l_∞ norm.

If the accumulation point x^* of the sequence $\{x_k\}$ generated from Algorithm 2.2 satisfies

$$Z^T \nabla f(x^*) = 0,$$

i.e., $\nabla f(x^*) \in N(Z^T)$, where $N(\cdot)$ denotes null space, then there is $\lambda^* \in R^m$ such that

$$\nabla f(x^*) = A\lambda^*,$$

which means any accumulation point x^* of the sequence $\{x_k\}$ generated from Algorithm 2.2 is a Kuhn-Tucker point of the original problem (1.1).

Next we give the global convergence theorem which says the reduced gradients converge to zero. Hence, any accumulation point of the sequence of iterates satisfies the first order necessary condition for a solution to (1.1).

Theorem 3.2 *Let $f : R^n \to R$ be continuously differentiable and bounded below on an feasible region. Assume $\{\hat{B}_k\}$ bounded uniformly, i.e., there is a positive constant M such that*

$$\|\hat{B}_k\| \leq M, \ \forall k.$$

Then, Algorithm 2.2 will terminate after finitely many iterations provided that $\{f(x_k), k = 1, 2, ...\}$ is bounded below . In other words, if $\epsilon = 0$, then either

$$\lim_{k\to\infty} f(x_k) = -\infty, \tag{3.8}$$

or

$$\liminf_{k\to\infty} \|\hat{g}_k\| = 0. \tag{3.9}$$

Proof. If the theorem is not true, then $f(x_k)$ is bounded below and there exists a positive constant δ such that

$$\|\hat{g}_k\| \geq \delta \tag{3.10}$$

which, togehter with Lemma 3.1, implies that

$$Pred_k(u_k) \geq \tau \min[1, \Delta_k] \tag{3.11}$$

for some positive constant τ.

Define the set

$$K_0 = \{k | \rho_k \geq \eta\}. \tag{3.12}$$

Inequality (3.11) and the assumption that $f(x_k)$ is bounded below give that

$$\sum_{k\in K_0} \Delta_k < \infty. \tag{3.13}$$

Because $\Delta_{k+1} \leq \xi_1 \Delta_k$ for all $k \notin K_0$, if follows from (3.13) that

$$\sum_{k=1}^{\infty} \Delta_k < \infty. \tag{3.14}$$

Therefore Therefore there exists $\bar{x}$ such that

$$\lim_{k\to\bar{x}} x_k = \bar{x}. \tag{3.15}$$

Relation (3.14) shows that $\Delta_k \to 0$. Thus it follows from (3.11) that

$$Pred_k(u_k) \geq \tau \Delta_k \tag{3.16}$$

for all sufficiently large k. (3.16) and the fact that $Pred_k = Ared_k + O(\Delta_k^2)$ indicate that

$$\lim_{k\to\infty} \rho_k = 1, \tag{3.17}$$

which yields that, for sufficiently large k,

$$\Delta_{k+1} \geq \Delta_k. \tag{3.18}$$

The above inequality contradicts (3.14). The contradiction proves the theorem. □

If $\mu > 0$, the convergence result can be further improved.

Theorem 3.3 *Under the conditions of Theorem 3.2, if $\mu > 0$, then every accumulation point of $\{x_k\}$ is a Kuhn-Tucker point of (1.1)-(1.2).*

Proof If the theorem is not true, there exist an accumulation point $\bar{x}^*$ which is not a KT point of problem Thus, there exist positive constants $\bar{\tau}$ and $\bar{\epsilon}$ such that

$$Pred_k(u_k) \geq \bar{\tau} \min[1, \Delta_k] \tag{3.19}$$

provided that $||x_k - \bar{x}^*|| \leq \bar{\epsilon}$. Define the sets

$$K_1 = \{k|\ \rho_k > \mu\} \tag{3.20}$$

$$\bar{K} = \{k|\ ||x_k - \bar{x}^*|| \leq \bar{\epsilon}\}. \tag{3.21}$$

Because $\mu > 0$, the set K_1 has similar properties as K_0 given in the proof of the previous theorem. Therefore it can be shown that

$$\sum_{k \in K_1 \cap \bar{K}} \Delta_k < \infty. \tag{3.22}$$

Hence there exists $\hat{k}$ such that

$$||x_{\hat{k}} - \bar{x}^*|| < \frac{1}{2}\bar{\epsilon}, \tag{3.23}$$

and

$$\sum_{k \in K_1 \cap \bar{K}, k \geq \hat{k}} \Delta_k < \frac{1}{2}\bar{\epsilon}. \tag{3.24}$$

The above two inequalities imply that $x_k \in \bar{K}$ for all $k \geq \hat{k}$. Therefore

$$\sum_{k=1}^{\infty} \Delta_k < \infty, \tag{3.25}$$

which implies that

$$\lim x_k = \bar{x}^*. \tag{3.26}$$

From the above relation, we can obtain a contradiction as in the proof of the previous theorem. This completes our proof. □

References

[1] K.A.Ariyawansa, Deriving collinear scaling algorithms as extensions of quasi-Newton methods and the local convergence of DFP and BFGS related collinear scaling algorithm, *Math. Prog.* 49 (1990) 23-48.

[2] R.H.Byrd, R.B.Schnabel and G.A.Shultz, A trust region algorithm for nonlinear constrained optimization, *SIAM J. Numer.Anal.* 24 (1987) 1152-1170.

[3] A.R.Conn, N.I.M.Gould and Ph.L.Toint, Global convergence of a class of trust region algorithms for optimization with simple bounds, *SIAM J. Numer. Anal.* 25 (1988) 433-460.

[4] Simon Di and Wenyu Sun, Trust region method for conic model to solve unconstrained optimization problems, to appear in *Optimization Methods and Software*.

[5] W.C.Davidon, Conic approximation and collinear scaling for optimizers, *SIAM J. Numer. Anal.* 17 (1980) 268-281.

[6] R.Fletcher, *Practical Methods of Optimization*, Second Edition, (John Wiley & Sons, 1987).

[7] Xuchu He and Wenyu Sun, *Introduction to Generalized Inverses of Matrices*, (Jiangsu Sci. & Tech. Publishing House, Nanjing, 1991.)

[8] J.J.Moré, Recent developments in algorithms and software for trust region methods, in *Mathematical Programming, The State of the Art, Bonn, 1982*, M.Grötschel and B.Korte eds., 258-287.

[9] J.J.Moré and D.C.Sorensen, Computing a trust region step, *SIAM J.Sci. Stat. Computing* 4 (1983) 553-572.

[10] J.L.Nazareth, Trust regions based on conic functions in linear and nonlinear programming, *J. Numerical Linear Algebra with Applications*, to appear.

[11] M.J.D.Powell, Convergence properties of a class of minimization algorithms, in *Nonlinear Programming 2*, O.L.Mangasarian, R.R.Meyer and S.M.Robinson, eds., (Academic Press, New York, 1975).

[12] M.J.D.Powell, On the global convergence of trust region algorithms for unconstrained minimization, *Mathematical Programming* 29 (1984) 297-303.

[13] M.J.D.Powell and Y.Yuan, A trust region algorithm for equality constrained optimization, *Mathematical Programming* 49 (1991) 189-211.

[14] Liqun Qi and Wenyu Sun, An iterative method for the minimax problem, in D.Z.Du and P.M.Pardalos eds., *Minimax and Applications*, (Kulwer Academic Publisher, Boston, 1995).

[15] Liqun Qi, Trust region algorithms for solving nonsmooth equations, *SIAM J. Optimization* 5 (1995) 219-230.

[16] D.C.Sorensen, The q-superlinear convergence of a collinear scaling algorithm for unconstrained optimization, *SIAM Journal of Numerical Analysis* 17 (1980) 84-114.

[17] D.C.Sorensen, Newton's method with a model trust region modification, *SIAM J. Numer. Anal.* 19 (1982) 409-429.

[18] Wenyu Sun, On nonquadratic model optimization methods, *Asia and Pacific Journal of Operations Research* No.4 (1995).

[19] Wenyu Sun and Zhonglin Wu, Optimization methods based on rational function model, *Journal of Mathematics* No.4 (1995).

[20] Wenyu Sun, On an adaptive approach for self-scaling variable metric algorithm, *Journal of Nanjing University, Mathematical Biquarterly* 10 (1993) 35-40.

[21] Wenyu Sun and Y.Yuan, *Optimization Theory and Methods*, (Science Press, Beijing, China, 1995).

[22] Y.Yuan, On a subproblem of trust region algorithms for constrained optimization, *Math. Prog.* 47 (1990) 53-63.