A GLOBALLY CONVERGENT METHOD OF CONSTRAINED MINIMIZATION BY SOLVING SUBPROBLEMS OF THE CONIC MODEL*

SUN LIN-PING [1)]
(*Department of Mathematics, Nanjing University, Nanjing, China*)

Abstract

A new method for nonlinearly constrained optimization problems is proposed. The method consists of two steps. In the first step, we get a search direction by the linearly constrained subproblems based on conic functions. In the second step, we use a differentiable penalty function, and regard it as the metric function of the problem. From this, a new approximate solution is obtained. The global convergence of the given method is also proved.

§1. Introduction

The nonlinearly constrained optimization problem to be considered in this paper is defined by

$$
\text{(NP)} \qquad \begin{array}{ll} \text{Minimize} & f(x), \\ \text{subject to} & e_i(x) \geq 0, \quad i = 1, 2, \cdots, m, \\ & h_j(x) = 0, \quad j = 1, 2, \cdots, l \end{array}
$$

where f, e_i, h_j denote real and differentiable functions of vector x in the n-dimensional Euclidean space $I\!R^n$.

Many techniques have been proposed to solve minimization problems with nonlinear constraints [2]. One of the proposed approaches is to iteratively solve linearly constrained subproblems. This method with quasi-Newton updates was originated by Han [4]. Powell [8] proposed another more practical update scheme, and proved that the methods having superlinear rates usually use a nondifferentiable penalty function, and regard it as the metric function of the problem [5]. Moreover, Yamashita [13] constructed a globally convergent

* Received February 18, 1986.

1) This paper was written while the author was visiting La Trobe University in Australia.

constrained quasi-Newton method, but his method is only suitable for the problems with inequality constraints. Murray and Wright 6[] studied the computation of the search direction in constrained optimization algorithms.

The above-mentioned methods are based on quadratic models. However, when the objective function has stronger non-quadratic properties in the neighbourhood of iterative points, these methods will face some difficulties. For this reason, we consider the methods based on the non-quadratic model. At present, conic models have been used in unconstrained minimization algorithms successfully (see Davidon [1], Gourgeon and Nocedal [3]). In this paper we establish a globally convergent method for nonlinearly constrained optimization problems. The method consists of two steps. In the first step, we get a search direction by solving linearly constrained subproblems based on the conic function. In the second step, we use a differentiable penalty function, and regard it as the metric function of the problem. From this, a new approximate solution is obtained. Section 2 gives the construction of the search direction. In Section 3 we establish the algorithm. In Section 4 we prove global convergence of the given method.

Except in Section 4, for convenience, in describing an iterative method we do not use superscripts to denote three neighbouring iterations containing the present iteration. Instead, we place a bar over or under quantities which correspond to the neighbouring iteration, e.g., if x denotes the present iteration, then $\bar{x}$ and $\underline{x}$ will denote the following and previous iteration, respectively. Subscripts are used to denote components of a vector, for example, x_i is the ith component of vector x.

§2 Construction of the Search Direction

In order to make a search direction d at iteration point x, we consider the subproblem that minimizes the conic function with linear constraints:

$$\text{(CCP)}\qquad \text{Minimize}\qquad c(x+d)=f(x)+\frac{\nabla f(x)^\tau d}{1+b^\tau d}+\frac{1}{2}\frac{d^\tau W d}{(1+b^\tau d)^2},$$

$$\begin{aligned}\text{subject to}\quad & e_i(x)+\nabla e_i(x)^\tau d\ge 0, \qquad i=1,\cdots,m,\\ & h_j(x)+\nabla h_j(x)^\tau d=0, \qquad j=1,\cdots,l,\\ & 1+b^\tau d>0\end{aligned}$$

where $c(x+d)$ is the approximation of $f(x)$ near x by the conic functon, $w=\nabla^2 f(x)+b\nabla f(x)^\tau+\nabla f(x)b^\tau$ and $b\in \mathbb{R}^n$.

Remark 2.1. The subproblem CCP is consistent. For example, $d=0$ is its special solution.

The solution of CCP and its corresponding Lagrange multiplier will be denoted by an array $(d,\sigma,\tau)\in \mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^l$ in the following discussion. By the Kuhn-Tucker condition of CCP we know that the array (d,σ,τ) must satisfy

$$\theta(I-\theta bd^\tau)(\nabla f(x)+\theta Wd)=\sum_{i=1}^{m}\sigma_i\nabla e_i(x)+\sum_{j=1}^{l}\tau_j\nabla h_j(x), \tag{2.1a}$$

$$\sigma_i(e_i(x)+\nabla e_i(x)d)=0,\quad i=1,\cdots,m, \tag{2.1b}$$

$$\theta_i\geq 0,\quad i=1,\cdots,m \tag{2.1c}$$

where $\theta=\dfrac{1}{1+b^\tau d}=1-\theta b^\tau d$.

On the other hand, in order to ensure global convergence of the method to be described later, we construct the metric function

$$F(x,u,v,\mu,\rho)=f(x)+\frac{1}{2}\sum_{i=1}^{m}\omega_i\mu_i(e_i(x)-\frac{u_i}{\mu_i})^2+\frac{1}{2}\sum_{j=1}^{l}\rho_j(h_j(x)-\frac{v_j}{\rho_j})^2 \tag{2.2}$$

where $0<u\in I\!R^m$ and $0<v\in I\!R^l$, $0<\mu\in I\!R^m$ and $0<\rho\in I\!R^l$ correspond to penalty parameter vectors, and

$$\omega_i=\begin{cases}1, & e_i(x)-u_i/\mu_i<0,\\ 0, & \text{otherwise.}\end{cases}$$

The following lemma will describe the descent property of the direction d.

Lemma 2.2. *If (d,σ,τ) is a solution of* CCP, *then inequality*

$$\nabla_x F(x,u,v,\mu,\rho)^\tau d\leq -\theta d^\tau Wd-\sum_{i=1}^{m}\varphi_i(x,u_i,\mu_i,\sigma_i)-\sum_{j=1}^{l}\psi_j(x,v_j,\rho_j,\tau_j) \tag{2.3}$$

holds, where

$$\varphi_i(x,u_i,\mu_i,\sigma_i)=(\omega_i\mu_i(e_i(x)-u_i/\mu_i)+\sigma_i/\theta^2)e_i(x), \tag{2.4}$$

$$\psi_j(x,v_j,\rho_j,\tau_j)=(\rho_j(h_j(x)-v_j/\rho_j)+\tau_j/\theta^2)h_j(x). \tag{2.5}$$

Proof. Since

$$\begin{aligned}\nabla_x F(x,u,v,\mu,\rho)^\tau d &= \nabla f(x)^\tau d+\sum_{i=1}^{m}\omega_i\mu_i(e_i(x)-u_i/\mu_i)\nabla e_i(x)^\tau d\\ &\quad+\sum_{j=1}^{l}\rho_j(h_j(x)-v_j/\rho_j)\nabla h_j(x)^\tau d\end{aligned}$$

we know by (2.1a)

$$\nabla f(x)^\tau d=\frac{1}{\theta^2}(\sum_{i=1}^{m}\sigma_i\nabla e_i(x)+\sum_{j=1}^{l}\tau_j\nabla h_j(x))^\tau d-\theta d^\tau Wd.$$

Thus

$$\begin{aligned}\nabla_x F(x,u,v,\mu,\rho)^\tau d &= -\theta d^\tau W d + \sum_{i=1}^{m}(\omega_i\mu_i(e_i(x)-u_i/\mu_i)+\sigma_i/\theta^2)\nabla e_i(x)^\tau d \\ &\quad + \sum_{j=1}^{l}(\rho_j(h_j(x)-v_i/\rho_j)+\tau_j/\theta^2)\nabla h_j(x)^\tau d.\end{aligned}$$

Moreover, since d is the solution of CCP, we obtain by (2.1b)

$$\begin{aligned}&\omega_i(e_i(x)-u_i/\mu_i)\nabla e_i(x)^\tau d \le -\omega_i(e_i(x)-u_i/\mu_i)e_i(x),\\ &\sigma_i\nabla e_i(x)^\tau d = -\sigma_i e_i(x),\\ &\nabla h_j(x)^\tau d = -h_j(x).\end{aligned}$$

Therefore

$$\begin{aligned}\nabla_x F(x,u,v,\mu,\rho)^\tau d &\le -\theta d^\tau W d - \sum_{i=1}^{m}(\omega_i\mu_i(e_i(x)-u_i/\mu_i)+\sigma_i/\theta^2)e_i(x) \\ &\quad - \sum_{j=1}^{l}(\rho_j(h_j(x)-v_j/\rho_j)+\tau_j/\theta^2)h_j(x).\end{aligned}$$

Lemma 2.2 shows that, if the parameters u_i, v_j, μ_i, ρ_j are adjusted properly, and W is positive definite, then the vector d can become a descent direction for the metric function F at x.

§3 Algorithm

Algorithm 3.1. For given $x, W, \varepsilon > 0$ and $r \in (0,1)$, we do as follows:

Step 1. Construct the subproblem CCP and find its Kuhn-Tucker array (d,σ,τ).

Step 2. If $\|d\| \le \varepsilon$, stop; or else go to Step 3.

Step 3. Compute u, v, μ and ρ.

Step 4. Select λ such that

$$F(x+\lambda d,u,v,\mu,\rho) \le F(x,u,v,\mu,\rho) + c_0\lambda\nabla_x F(x,u,v,\mu,\rho)^\tau d$$

where $c_0 = 0.0001$.

Step 5. Set $\bar{x} = x + \lambda d$.

Step 6. Calculate $\bar{b}$.

Step 7. Update $\bar{B}$ by some scheme.

Step 8. Calculate $\bar{W} = \bar{B} + \bar{b}\nabla f(\bar{x})^\tau + \nabla f(\bar{x})\bar{b}^\tau$.

Step 9. Set $x = \bar{x}, W = \bar{W}, b = \bar{b}$ and go to Step 1.

Remark 3.2. (a) A numerical method for solving the subproblem CCP was proposed by Sun [11].

(b) Step 3 is a procedure for finding the approximate values of u, v, μ and ρ such that d becomes a descent direction for the function $F(x,u,v,m,\rho)$ at x. We usually compute by the following algorithm:

Algorithm 3.3. It is given that $\eta_1 \in (0,1), \underline{u}, \underline{v}, \underline{\mu}, \underline{\rho}, \Delta u_i > 0, \Delta v_j > 0$ and $r > 1$.

Step 1. Set $u = \underline{u}, v = \underline{v}, \mu = \underline{\mu}$ and $\rho = \underline{\rho}$.

Step 2. If

$$\sum_{i=1}^{m} \varphi_i(xu_i, \mu_i\sigma_i) + \sum_{j=1}^{l} \psi_j(x, v_j, \rho_j, \tau_j) \geq -\eta_1 \theta d^T W d \tag{3.1}$$

then stop, or else go to Step 3.

Step 3. Define

$$\begin{aligned} I &= \{i \mid \varphi_i(x, u_i, \mu_i, \sigma_i) < 0, \quad i = 1, \cdots, m\}, \\ I_1 &= \{i \mid e_i(x) < 0, i \in I\}, \\ I_2 &= \{i \mid 0 < e_i(x) < u_i/\mu_i, i \in I\}, \\ J &= \{j \mid \psi_j(x, v_j, \rho_j, \tau_j) < 0, \quad j = 1, \cdots, l\}, \\ J_1 &= \{j \mid h_j(x) < 0, j \in J\}, \\ J_2 &= \{j \mid 0 < h_j(x) < v_j/\rho_j, j \in J\}. \end{aligned}$$

Step 4. Set

$$\begin{aligned} u_i &= u_i + \Delta u_i, \quad i \in I_1, \\ u_i &= \max\{r\mu_i, u_i/e_i(x)\}, \quad i \in I_2, \\ v_j &= v_j + \Delta v_j, \quad j \in J_1, \\ \rho_j &= \max\{r\rho_j, v_j/h_j(x)\}, \quad j \in J_2. \end{aligned}$$

and go to Step 2.

Algorithm 3.3 shows that, if necessary, u_i and v_j are increased when $e_i(x) < 0$ and $h_j(x) < 0$ respectively, and that μ_i and ρ_j are increased when $0 < e_i(x) < u_i/\mu_i$ and $0 < h_j(x) < v_j/\rho_j$ respectively. Therefore, if W is positive definite, it is known by a simple calculation that Algorithm 3.3 has the property of finite termination.

(c) b is obtained by the following procedure. From CCP, we know that the conic function $c(x+d)$ satisfies

$$c(x) = f(x), \nabla c(x) = \nabla f(x), \nabla^2 c(x) = \nabla^2 f(x). \tag{3.2}$$

Let κ be the number of iteration times. If we impose the requirement

$$c(\underline{x}^{(t)}) = f(\underline{x}^{(t)}), \quad t = 1, \cdots, \kappa_p \tag{3.3}$$

for past iterations $\underline{x}^{(t)}$ and $\kappa_p = \min\{\kappa - 1, n\}$, then

$$b^T s^{(t)} = q_t \tag{3.4}$$

where

$$s^{(t)} = x - x^{(t)}, \tag{3.5}$$

$$q_t = 1 + \frac{a_1 - \sqrt{a_1^2 + a_2(a + a_1)}}{a_2} \tag{3.6}$$

where

$$a = \frac{1}{2}s^{(t)}Bs^{(t)}, \quad a_1 = \nabla f(x)^\tau s^{(t)}, \quad a_2 = f(\underline{x}^{(t)}) - f(x)$$

(see Schnabel [10]). In this case, b is defined by the linear equations

$$Sb = q \tag{3.7}$$

that is,

$$b = S^+ q \tag{3.8}$$

where S is constructed by $s^{(t)}$ as row vectors, and S^+ is the generalised inverse matrix of S. We can obtain S^+ readily by using Greville's algorithm with successive row recursion (see Sun [12]).

(d) Updating B is most effectively done by the quasi-Newton methods. The matrix B is intended to be an approximation of the Hessian of the Lagrange function

$$L(x, \sigma, \tau) = f(x) + \sum_{i=1}^{m} \sigma_i e_i(x) + \sum_{j=1}^{l} \tau_j h_j(x). \tag{3.9}$$

We follow Powell's recommendation [8]. Thus B is updated by the BFGS formula:

$$\bar{B} = B + \frac{\delta\delta^\tau}{S^\tau\delta} - \frac{Bss^\tau B}{s^\tau Bs} \tag{3.10}$$

where $s = \bar{x} - x$, $\delta = \eta_2 y + (1 - \eta_2)Bs$, $y = \nabla_x L(\bar{x}, \sigma, \tau) - \nabla_x L(x, \sigma, \tau)$ and $\eta_2 \in (0, 1)$ is chosen such that $s^\tau\delta > 0$.

§4 Global Convergence

Now we discuss the global convergence of Algorithm 3.1.

Lemma 4.1. *The functions $\mu_i(e_i(x) - u_i/\mu_i)^2, i = 1, \cdots, m$, are monotone decreasing functions of μ_i when μ_i satisfies $|e_i(x)| \le u_i/\mu_i$.*

Proof. Because

$$\frac{\partial(\mu_i(e_i(x) - u_i/\mu_i)^2)}{\partial\mu_i} = (e_i(x) - u_i/\mu_i)(e_i(x) + u_i/\mu_i) \tag{4.1}$$

the formula (4.1) is not positive when $|e_i(x)| \le u_i/\mu_i$.

We have the following lemma that is similar to Lemma 4.1.

Lemma 4.2. *The functions $\rho_j(h_j(x)-v_j/\rho_j)^2, j=1,\cdots,l$, are monotone decreasing functions of ρ_j satisfying $|h_j(x)|\le v_j/\rho_j$.*

The proof of this lemma is like that of Lemma 4.1.

For clarity, we use superscripts to denote the κth iteration of the algorithm in the following statements.

Lemma 4.3. *Let the Lagrange multipliers $\sigma^{(\kappa)}$ and $\tau^{(\kappa)}$ of CCP be bounded; then there exists an integer $\kappa_0\ge 0$ such that $u^{(\kappa)}=u^{(\kappa_0)}$ and $v^{(\kappa)}=v^{(\kappa_0)}$ for $\kappa\ge\kappa_0$.*

Proof. Suppose that the conclusion of the lemma is false. Then there is an index i such that $u_i^{(\kappa)}\to\infty$, because $u_i^{(\kappa)}$ increases monotonically with a fixed step-length $\Delta u_i>0$.

On the other hand, we have $\varphi_i(x,u_i,\mu_i,\sigma_i)\ge 0$ if $u_i^{(\kappa)}\ge\sigma_i^{(\kappa)}/\theta_\kappa^2$ and $e_i(x^{(\kappa)})<0$. Thus $u_i^{(\kappa)}\to\infty$ is impossible by the fact that $u_i^{(\kappa)}$ is increased only when $e_i(x^{(\kappa)})<0$ and that $\sigma_i^{(\kappa)}$ is bounded. In other words, there is an integer $\kappa_1\ge 0$ such that $u^{(\kappa)}=u^{(\kappa_1)}$ for $\kappa\ge\kappa_1$. For the same reason we also know there is an integer $\kappa_2\ge 0$ such that $v^{(\kappa)}=v^{(\kappa_2)}$ for $\kappa\ge\kappa_2$. Set $\kappa_0=\max\{\kappa_1,\kappa_2\}$; then the lemma is proved.

A principal result of this paper is stated as follows:

Theorem 4.4. *In the problem* (NP), *assume that the following conditions are satisfied:*

(a) *There are two positive numbers $\alpha_1>0$ and $\alpha_2>0$ such that*

$$\alpha_1 z^\tau z\le z^\tau W_\kappa z\le a_2 z^\tau z,\quad \forall z\in I\!R^n.$$

(b) *There is a compact set $\Omega\in I\!R^n$ such that $x^{(\kappa)}\in\Omega$.*

(c) *There is a Kuhn-Tucker array $(d^{(\kappa)},\sigma^{(\kappa)},\tau^{(\kappa)})$ of CCP, and this array is bounded. Then any infinite sequence $\{x^{(\kappa)}\}$ generated by Algorithm 3.1 has the property*

$$\lim_{\kappa\to\infty}\inf\|d^{(\kappa)}\|=0.$$

Proof. From Lemma 2.1 and inequality (3.1) we have

$$\nabla_x F(x^{(\kappa)},u^{(\kappa)},v^{(\kappa)},\mu^{(\kappa)},\rho^{(\kappa)})^\tau d^{(\kappa)}\le-(1-\eta_1)\theta_k(d^{(\kappa)})^\tau W_\kappa d^{(\kappa)}.$$

Thus, from Step 4 of Algorithm 3.1 we get

$$\begin{aligned}F(x^{(\kappa+1)},\ &u^{(k)},v^{(\kappa)},\mu^{(\kappa)},\rho^{(k)})-F(x^{(\kappa)},u^{(\kappa)},v^{(\kappa)},\mu^{(\kappa)},\rho^{(\kappa)})\\ &\le-c_0\lambda_\kappa(1-\eta_1)\theta_\kappa(d^{(\kappa)})^\tau W_\kappa d^{(\kappa)}.\end{aligned}$$

Moreover, by Lemma 4.1, Lemma 4.2 and Algorithm 3.3 we have

$$F(x^{(\kappa+1)},u^{(\kappa)},v^{(\kappa)},\mu^{(k+1)},\rho^{(\kappa+1)})\le F(x^{(\kappa+1)},u^{(\kappa)},v^{(\kappa)},\mu^{(\kappa)},\rho^{(\kappa)}).$$

From Lemma 4.3 we get

$$\begin{aligned}F(x^{(\kappa+1)},u^{(\kappa+1)},\ &v^{(\kappa+1)},\mu^{(\kappa+1)},\rho^{(\kappa+1)})-F(x^{(\kappa)},u^{(\kappa)},v^{(\kappa)},\mu^{(\kappa)},\rho^{(\kappa)})\\ &\le-c_0\lambda_\kappa(1-\eta_1)\theta_\kappa(d^{(\kappa)})^\tau W_\kappa d^{(\kappa)}\end{aligned}\tag{4.2}$$

for sufficiently large κ.

Suppose that the conclusion of the theorem is false. If $\{\mu^{(\kappa)}\} < \infty$ and $\{\rho^{(\kappa)}\} < \infty$ we can obtain the desired conclusion by the usual convergence proof of Armij's step-length rule. Therefore we consider the following situation, i.e. when $\mu^{(\kappa)} \to \infty$ and $\rho^{(\kappa)} \to \infty$, there exists an $\varepsilon_0 > 0$ such that $\|d^{(\kappa)}\| \geq \varepsilon_0$.

Since $F(x,u,v,\mu,\rho)$ is bounded below by $f(x)$, by assumption (b), and as the sequence $(x^{(\kappa)}, u^{(\kappa)}, v^{(\kappa)}, \mu^{(\kappa)}, \rho^{(\kappa)})$ is bounded below, from (4.2) and assumption (a) we have

$$\infty > \sum_{\kappa} \lambda_\kappa \theta_\kappa (d^{(\kappa)})^\tau W_\kappa d^{(\kappa)} \geq \alpha_1 \sum_{\kappa} \lambda_\kappa \theta_\kappa \|d^{(\kappa)}\|^2.$$

Let $k \triangleq \{\kappa | \kappa \geq \kappa_0, \mu^{(\kappa)} \neq \mu^{(\kappa-1)}$ and $\rho^{(\kappa)} \neq \rho^{(\kappa-1)}\}$, we may assume, without loss of generality, that it is an infinite set. From Algorithm 3.3 we get

$$\sum_{i=1}^{m} \varphi_i(x^{(\kappa)}, u_i^{(\kappa-1)}, \mu_i^{(\kappa-1)}, \sigma_i^{(\kappa)}) + \sum_{j=1}^{l} \psi_j(x^{(\kappa)}, v_j^{(\kappa-1)}, \rho_j^{(\kappa-1)}, \tau_j^{(\kappa)}) < -\eta_1 \theta_\kappa (d^{(\kappa)})^\tau W_\kappa d^{(\kappa)}. \tag{4.3}$$

Define

$$i_0^{(\kappa)} = \arg\min\{\varphi_i(x^{(\kappa)}, u_i^{(\kappa-1)}, \mu_i^{(\kappa-1)}, \sigma_i^{(\kappa)}) | \quad i = 1, \cdots, m\},$$

$$j_0^{(\kappa)} = \arg\min\{\psi_j(x^{(\kappa)}, v_j^{(\kappa-1)}, \rho_j^{(\kappa-1)}, \tau_j^{(\kappa)}) | \quad j = 1, \cdots, l\}.$$

Thus

$$\begin{aligned} & m\varphi_{i_0}(x^{(\kappa)}, u_{i_0}^{(\kappa-1)}, \mu_{i_0}^{(\kappa-1)}, \sigma_{i_0}^{(\kappa)}) + l\psi_{j_0}(x^{(\kappa)}, v_{j_0}^{(\kappa-1)}, \rho_{j_0}^{(\kappa-1)}, \tau_{j_0}^{(\kappa)}) \\ & \leq \sum_{i=1}^{m} \varphi_i(x^{(\kappa)}, u_i^{(\kappa-1)}, \mu_i^{(\kappa-1)}, \sigma_i^{(\kappa)}) + \sum_{j=1}^{l} \psi_j(x^{(\kappa)}, v_j^{(\kappa-1)}, \rho_j^{(\kappa-1)}, \tau_j^{(\kappa)}) \\ & < -\eta_1 \theta_\kappa (d^{(\kappa)})^\tau W_\kappa d^{(\kappa)}. \end{aligned}$$

Notice that $\mu_{i_0}^{(\kappa)} \to \infty$ and $\rho_{j_0}^{(\kappa)} \to \infty$ by the definition of i_0. By taking the relations

$$0 < e_{i_0}(x^{(\kappa)}) < u_{i_0}^{(\kappa-1)} / \mu_{i_0}^{(\kappa-1)},$$

$$0 < h_{j_0}(x^{(\kappa)}) < v_{j_0}^{(\kappa-1)} / \rho_{j_0}^{(\kappa-1)}$$

into account, we have respectively

$$\begin{aligned} \varphi_{i_0}(x^{(\kappa)}, u_{i_0}^{(\kappa-1)}, \mu_{i_0}^{(\kappa-1)}, \sigma_{i_0}^{(\kappa)}) &= \omega_{i_0}^{(\kappa)} \mu_{i_0}^{(\kappa-1)} (e_{i_0}(x^{(\kappa)}))^2 + (\sigma_{i_0}^{(\kappa)}/\theta_\kappa^2 - \omega_{i_0}^{(\kappa)} u_{i_0}^{(\kappa-1)}) e_{i_0}(x^{(\kappa)}) \\ &> -(|\sigma_{i_0}^{(\kappa)}/\theta_\kappa^2 - \omega_{i_0}^{(\kappa)} u_{i_0}^{(\kappa-1)}| u_{i_0}^{(\kappa-1)}) / \mu_{i_0}^{(\kappa-1)} \geq -\beta_1 / \mu_{i_0}^{(\kappa-1)} \end{aligned}$$

and

$$\begin{aligned} \psi_{j_0}(x^{(\kappa)}, v_{j_0}^{(\kappa-1)}, \rho_{j_0}^{(\kappa-1)}, \tau_{j_0}^{(\kappa)}) &= \rho_{j_0}^{(\kappa-1)} (h_{j_0}(x^{(\kappa)}))^2 + (\tau_{j_0}^{(\kappa)}/\theta_\kappa^2 - v_{j_0}^{(\kappa-1)}) h_{j_0}(x^{(\kappa)}) \\ &> -(|\tau_{j_0}^{(\kappa)}/\theta_\kappa^2 - v_{j_0}^{(\kappa-1)}| v_{j_0}^{(\kappa-1)}) / \rho_{j_0}^{(\kappa-1)} \geq \beta_2 / \rho_{j_0}^{(\kappa-1)} \end{aligned}$$

where β_1 and β_2 are constants satisfying

$$\beta_1 \geq |\sigma_{i_0}^{(\kappa)}/\theta_\kappa^2 - \omega_{i_0}^{(\kappa)} u_{i_0}^{(\kappa-1)}|u_{i_0}^{(\kappa-1)} \text{ and } \beta_2 \geq |\tau_{j_0}^{(\kappa)}/\theta_\kappa^2 - v_{j_0}^{(\kappa-1)}|v_{j_0}^{(\kappa-1)}.$$

Thus

$$m\beta_1/\mu_{j_0}^{(\kappa-1)} + l\beta_2/\rho_{j_0}^{(\kappa-1)} > \eta_1\theta_\kappa\alpha_1\|d^{(\kappa)}\|^2, \quad \kappa \in k.$$

The above inequality shows a contradiction because $\mu_{i_0}^{(\kappa-1)} \to \infty$, $\rho_{j_0}^{(\kappa-1)} \to \infty$, and at the same time $\|d^{(\kappa)}\| \geq \varepsilon_0 > 0$. Thus the theorem is proved.

This theorem shows that if $\varepsilon = 0$, the algorithm either terminates at a Kuhn-Tucker point of CCP or generates an infinite sequence $\{x^{(\kappa)}\}$ in which there exists a subsequence $\{x^{(\kappa)}\}, \kappa \in k$, such that $d^{(\kappa)} \to 0$. Also, the theorem shows that for any given $\varepsilon > 0$, there is a finite $\kappa(\varepsilon)$ such that we have $\|d^{(\kappa)}\| < \varepsilon$.

By Theorem 4.4 we obtain directly the following theorem describing convergence.

Theorem 4.5. *Let the assumptions of Theorem* 4.4 *hold; then any infinite sequence* $\{x^{(\kappa)}\}$ *generated by Algorithm* 3.1 *has at least one accumulation point* $x^* \in \mathbb{R}^n$ *such that* x^* *is a* Kuhn-Tucker *point of* (NP).

Proof. From Theorem 4.4 we know that there is a subsequence $\{x^{(\kappa)}\}, \kappa \in k$, such that $d^{(\kappa)} \to 0$. Without loss of generality we may assume that there are $x^* \in \Omega, \sigma^* \in \mathbb{R}^m, \tau^* \in \mathbb{R}^l$ and positive definite $W^* \in \mathbb{R}^{n\times n}$ such that

$$x^{(\kappa)} \to x^*, \quad \sigma^{(\kappa)} \to \sigma^*, \quad \tau^{(\kappa)} \to \tau^* \text{ and } W_\kappa \to W^*$$

where $\kappa \in k$. In (2.1) let $\kappa \to \infty, \kappa \in k$. Since $\theta_\kappa \to 1$, we have

$$\nabla f(x^*) - \sum_{i=1}^m \sigma_i^* \nabla e_i(x^*) - \sum_{j=1}^l \tau_j^* \nabla h_j(x^*) = 0,$$
$$\sigma_i^* e_i(x^*) = 0, \quad i = 1, \cdots, m,$$
$$\sigma_i^* \geq 0, \quad i = 1, \cdots, m$$

and

$$e_i(x^*) \geq 0, \quad i = 1, \cdots, m,$$
$$h_j(x^*) = 0, \quad j = 1, \cdots, l.$$

Thus x^* is a Kuhn-Tucker point of (NP).

References

[1] W.C. Davidon, Conic approximations and collinear scalings for optimizers, *SIAM J. Num. Anal.*, **17** (1980), 268–281.

[2] P.E. Gill, W. Murray, Practical Optimization, Academic Press, New York, 1981.

[3] H. Gourgeon, J. Nocedal, A conic algorithm for optimization, *SIAM J. Sci. Stat. Comput.*, 6 (1985), 255-267.

[4] S.P. Han, Superlinearly convergent variable metric algorithms for general nonlinear programming problems, *Math. Prog.*, 11 (1976), 263–282.

[5] S.P. Han, A globally convergent method for nonlinear programming, *JOTA* 22, 1977, 297–309.

[6] W. Murray, M.H. Wright, Computation of the search direction in constrained optimization algorithms, *Math. Prog. Study*, 16 (1982), 45-61.

[7] J.M. Ortega, M.C. Rheinbolt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[8] M.J.D. Powell, A fast algorithm for nonlinearly constrained optimization calculations, in : G.A. Watson, ed., Num. Anal., 1978, 144–157.

[9] M.J.D. Powell, The convergence of variable metric methods for nonlinearly constrained optimization calculation, in : O.L. Mangasarian, R.R. Meyer and S.M. Robinson, eds., Nonlinear Programming 3, 1978, 27–63.

[10] R.B. Schnabel, Unconstrained minimization using conic models and second derivatives, Dept. Comput. Sci., Report Cu-227-82, Univ. Colorado, Boulder, 1982.

[11] Sun Linping, A method that minimizes conical functions with linear constraints, 1985. (to appear)

[12] Sun Linping , An unconstrained minimization method based on nonquadratic model, *J. Num. Math. Comp. Appl.*, 6 (1985), 188–193.

[13] H. Yamashita, A globally convergent constrained quasi-Newton method with an augmented Lagrangian type penalty function, *Math. Prog.*, 23 (1982), 75–86.