Journal of Computational Mathematics, Vol.22, No.5, 2004, 651–660.

A QP FREE FEASIBLE METHOD [*1)]

Ding-guo Pu Yan Zhou Hai-yan Zhang
(*Department of Mathematics, Tongji University, Shanghai 200092, China*)

Abstract

In [12], a QP free feasible method was proposed for the minimization of a smooth function subject to smooth inequality constraints. This method is based on the solutions of linear systems of equations, the reformulation of the KKT optimality conditions by using the Fischer-Burmeister NCP function. This method ensures the feasibility of all iterations. In this paper, we modify the method in [12] slightly to obtain the local convergence under some weaker conditions. In particular, this method is implementable and globally convergent without assuming the linear independence of the gradients of active constrained functions and the uniformly positive definiteness of the submatrix obtained by the Newton or Quasi Newton methods. We also prove that the method has superlinear convergence rate under some mild conditions. Some preliminary numerical results indicate that this new QP free feasible method is quite promising.

Mathematics subject classification: 90C30, 65K10.
Key words: Constrained optimization, KKT point, Multiplier, Nonlinear complementarity, Convergence.

1. Introduction

Consider the constrained nonlinear optimization Problem (NLP):

$$\min f(x), \quad x \in R^n, \quad \text{s.t.G(x)} \leq 0,$$

where $f : R^n \to R$ and $G(x) = (g_1(x), g_2(x), \cdots, g_m(x))^T : R^n \to R^m$ are Lipchitz continuously differentiable functions.

We denote by $D = \{x \in R^n | G(x) < 0\}$ and $\bar{D} = cl(D)$ the strictly feasible set and the feasible set of Problem (NLP), respectively.

The Lagrangian function associated with Problem (NLP) is the function

$$L(x, \lambda) = f(x) + \lambda^T G(x), \tag{1}$$

where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_m)^T \in R^m$ is the multiplier vector. For simplicity, we use (x, λ) to denote the column vector $(x^T, \lambda^T)^T$.

A *Karush-Kuhn-Tucker* (KKT) point $(\bar{x}, \bar{\lambda}) \in R^n \times R^m$ is a point that satisfies the necessary optimality conditions for Problem (NLP):

$$\nabla_x L(\bar{x}, \bar{\lambda}) = 0, \ G(\bar{x}) \leq 0, \ \bar{\lambda} \geq 0, \ \bar{\lambda}_i g_i(\bar{x}) = 0, \tag{2}$$

where $1 \leq i \leq m$. We also say $\bar{x}$ is a KKT point if there exists a $\bar{\lambda}$ such that $(\bar{x}, \bar{\lambda})$ satisfy (2).

Finding KKT points for Problem (NLP) can be equivalently reformulated as solving the mixed nonlinear complementarity problem (NCP) in (2), Problem (NCP) has attracted much

* Received April 4, 2002; final revised March 10, 2004.
1) This research is supported by National Science Foundation of China (No.10371089).

attention due to its various applications, see [4, 1]. One method to solve the nonlinear complementarity problem is to construct a Newton method for solving a system of nonlinear equations (see [11, 5]). Qi and Qi [12] proposed a new QP-free method which ensures the strict feasibility of all iterates. Their work is based on the Fischer-Burmeister NCP function. They proved the global convergence without isolatedness of the accumulation point and the strict complementarity condition. They also proved the superlinear convergence under mild conditions.

However, for the global convergence, [12] still used some stronger conditions. One is the linear independence of the gradients of active constrained functions at the solution; another is the uniformly positive definiteness of H^k which is obtained by the quasi Newton update. To overcome the shortcoming, in this paper, an algorithm is proposed for the minimization of a smooth function subject to smooth inequality constraints. This algorithm is based on the method in [12]. Our main work is to modify this method slightly for obtaining the global convergence under some weaker conditions. Comparing with the method in [12], our method is implementable and globally convergent without assuming the uniformly positive definiteness of H^k and the linear independence of the gradients of active constrained functions at the solution. In particular, for the superlinear convergence of the algorithm we used the same conditions as the method in [12].

In this paper, we use the Fischer-Burmeister function [2] as the following:

$$\psi(a,b) = \sqrt{a^2+b^2} - a - b.$$

Let $\phi_i(x,\lambda) = \psi((-g_i(x)),\lambda_i)$, $1 \le i \le m$, $\Phi_1(x,\lambda) = (\phi_1(x,\lambda)\cdots\phi_m(x,\lambda))^T$. We denote $\Phi(x,\lambda) = ((\nabla_x L(x,\lambda))^T, (\Phi_1(x,\lambda))^T)^T$, Clearly, the KKT point conditions of (2) are equivalently reformulated as the condition $\Phi(x,\lambda) = 0$.

Let $I_1(x,\lambda) = \{i|(g_i(x),\lambda_i) \ne (0,0)\}$ and $I_0(x,\lambda) = \{i|(g_i(x),\lambda_i) = (0,0)\}$. If $j \in I_1(x,\lambda)$, then we denote,

$$\xi_j(x,\lambda) = \frac{g_j}{\sqrt{(g_j)^2+(\lambda_j)^2}} + 1; \qquad \gamma_j(x,\lambda) = \frac{\lambda_j}{\sqrt{(g_j)^2+(\lambda_j)^2}} - 1.$$

We have $\nabla_x\phi_j = \xi_j \nabla g_j(x)$ and $\nabla_\lambda \phi_j = \gamma_j e_j$ where $e_j = (0,\cdots,0,1,0\cdots,0)^T \in R^m$ is the jth column of the unit matrix, its jth element is 1, and other elements are 0. If $j \in I_0(x,\lambda)$, then we denote

$$\xi_j(x,\lambda) = 1 - \sqrt{2}/2; \qquad \gamma_j = \gamma_j(x,\lambda) = -1 + \sqrt{2}/2.$$

We have $\xi_j \nabla g_j(x) \in \partial_x \phi_j(x,\lambda)$ and $\gamma_j e_j \in \partial_\lambda \phi_j(x,\lambda)$. Clearly, $\xi_j^2 + \gamma_j^2 \ge 3 - 2\sqrt{2} > 0$.

The paper is organized as follows. In Section 2, we propose a QP free feasible method. In Section 3, we show that the algorithm is well defined. In Section 4 and Section 5, we discuss the conditions of the global convergence and superlinear convergence of the algorithm, respectively. In Section 6, we give a brief discussion and some numerical tests.

2. Algorithm

In the following algorithm 2.1, let $\xi_j^k = \xi_j(x^k,\mu^k)$ and $\gamma_j^k = \gamma_j(x^k,\mu^k)$, $\eta_j^k = -\sqrt{-2\gamma_j^k}$,

$$V^k = \begin{pmatrix} V_{11}^k & V_{12}^k \\ V_{21}^k & V_{22}^k \end{pmatrix} = \begin{pmatrix} H^k + \bar{c}_1^k I_n & \nabla G^k \\ diag(\xi^k)(\nabla G^k)^T & diag(\eta^k - c^k) \end{pmatrix},$$

where I_n is the n order unit matrix, $\bar{c}_1^k = c_1 min\{1, \|\bar{\Phi}^k\|^\nu\}$, $\bar{\Phi}^k = \Phi(x^k,\bar{\lambda}^k)$, $\bar{\lambda}^k$ is obtained in Algorithm 2.1, $c_1 \in (0,1)$, $diag(\xi^k)$ or $diag(\eta^k - c^k)$ denotes the diagonal matrix whose jth