Journal of Computational Mathematics, Vol.21, No.2, 2003, 247–256.

A ROBUST SQP METHOD FOR OPTIMIZATION WITH INEQUALITY CONSTRAINTS[*1)]

Juliang Zhang[†]
(*School of Economics and Manggement, Tsinghua University, Beijing, 100084, China*)
Xiangsun Zhang
(*School of Information Engineering, Beijing University of Posts and Telecommunications, Beijing, 100876, China*)

Abstract

A new algorithm for inequality constrained optimization is presented, which solves a linear programming subproblem and a quadratic subproblem at each iteration. The algorithm can circumvent the difficulties associated with the possible inconsistency of QP subproblem of the original SQP method. Moreover, the algorithm can converge to a point which satisfies a certain first-order necessary condition even if the original problem is itself infeasible. Under certain condition, some global convergence results are proved and local superlinear convergence results are also obtained. Preliminary numerical results are reported.

Key words: nonlinear optimization, SQP method, global convergence, superlinear convergence.

1. Introduction

We consider the following nonlinear programming problem:

$$\begin{array}{ll} \min\limits_{x\in R^n} & f(x) \\ \text{s.t.} & c_i(x)\le 0, \quad i\in I \end{array} \tag{1}$$

where $f: R^n \to R$, $c_i : R^n \to R$, $i \in I$ are continuously differentiable functions. $I = \{1, 2, \ldots, m\}$. Let $g(x) = \nabla f(x)$, $C(x) = (c_1(x), c_2(x), \ldots, c_m(x))^T$ and $A(x) = (\nabla c_1(x), \nabla c_2(x), \ldots, \nabla c_m(x))$. In view of convenience, we usually use f_k for $f(x_k)$, C_k for $C(x_k)$, g_k for $g(x_k)$ and A_k for $A(x_k)$, etc.

SQP algorithms for constrained optimization are iteration-type methods. They generate a sequence of points approximating to the solution by the procedure

$$x_{k+1} = x_k + \lambda_k d_k \tag{2}$$

where x_k is the current point, d_k is a search direction which minimizes a quadratic model subject to linearized constraints and λ_k is the stepsize along the search direction (see details in [12, 18, 25]). For $k \ge 1$ the original SQP method developed by Wilson, Han and Powell employs the following QP subproblem

$$\begin{array}{ll} \min & g_k^T d + \frac{1}{2} d^T B_k d \\ \text{s.t.} & C_k + A_k^T d \le 0 \end{array} \tag{3}$$

where B_k is a symmetric matrix which approximates to the $Hessian$ of the $Lagrangian$ function

$$L(x,\rho) = f(x) + \rho^T C(x) \tag{4}$$

---

* Received December 1, 2000.
1)This work is supported in part by the National Natural Science Foundation of China (Grant No. 10171055).
† Email address: jlzhang@mail.amt.ac.cn.

where ρ is an approximation to the Lagrangian multiplier vector.

Because of its nice convergence properties(for example, see Han (1977), Powell (1978), Bogg et al (1982)), the SQP method has been absorbing attention from many researchers.

The requisite consistency of the linearized constraints of the QP subproblem (3) is a serious limitation of the SQP method. Within the framework of the method, Powell suggested to solve a modified subproblem at each iteration (Powell(1977)):

$$\begin{array}{ll} \min & g_k^T d + \frac{1}{2} d^T B_k d + \frac{1}{2}\delta_k (1-\mu)^2 \\ \text{s.t.} & \mu_i c_i(x_k) + \nabla c_i(x_k)^T d \le 0 \end{array} \tag{5}$$

where $\mu_i = \begin{cases} 1, & c_i(x_k) < 0 \\ \mu, & c_i(x_k) \ge 0 \end{cases}$ and $0 \le \mu \le 1$, $\delta_k > 0$ is a penalty parameter. With some other technique, the computational investigation provided by Schittkowski (1981,1983) shows that this modification works very well.

However, a simple example presented by Burke and Han (1989) and Burke (1992) indicates that this approach may not be the best one.

On the other hand, based on the trust region strategy, Fletcher(1981) developed the Sl_1QP method for problem (1). Burke and Han (1989) shows that Fletcher's approach is still incomplete. One of the reasons is that the search direction generated by Sl_1QP method may point to the contrary of the optimal point.

Burke and Han (1989) and Burke (1989) presented approaches to overcome difficulties associated with the inconsistency of the QP subproblem (3). A feature different to the other methods is that even when (1) is itself infeasible their methods can converge to a point which meets a certain first-order necessary optimality condition. However, Burke and Han's method is conceptual.

Recently, Liu and Yuan (2000) presented a method which is a modification to SQP method. Similar to Burke and Han's methods, even when (1) is itself infeasible their method can converge to a point which meets a certain first-order necessary optimality condition. Unlike the other methods, their method solves two subproblem—one is an unconstrained piecewise quadratic subproblem, the other is a quadratic subproblem. Their method has excellent theoretical properties and is implementable.

In this paper, we describe another implementable method which is a modification to SQP method. The algorithm can circumvent difficulties associated with the infeasibility of the QP subproblem. Our method is similar to Liu and Yuan's method. At each iteration it solves two subproblems—one is a linear programming, which is different from Liu and Yuan's, the other is a quadratic subproblem. Since solving a linear programming is much easier than solving a piecewise quadratic programming, the computation at each iteration is less than Liu and Yuan's. Under certain conditions we can prove that the method is globally convergent and locally superlinearly convergent.

Note that in our method we only deal with inequality constraints. In fact, for equality constraints, we can convert it into two inequalities and our method can also deal with it. Therefore, our method can solve optimization problem with general constraints.

Our algorithm can be easily combined with the trust region strategy. Thus the algorithm in this paper can be extended to a trust region algorithm for constrained optimization problem.

The paper is organized as follows. The algorithm model is presented in Section 2. In Section 3, the global convergence results of the algorithm are proved. We discuss the local properties of the algorithm in Section 4. In Section 5, some numerical results are reported. Some discussions are given in Section 6 to conclude the paper.