International Journal of Pure and Applied Mathematics
Volume 61 No. 2 2010, 161-167

A VARIANT OF FEASIBLE DESCENT SQP METHOD FOR INEQUALITY CONSTRAINED OPTIMIZATION

Zhijun Luo[1 §], Guohua Chen[2], Jinglong Liang[3]
[1,2,3]Department of Mathematics and Applied Mathematics
Hunan Institute of Humanities, Science and Technology
Loudi, 417000, P.R. CHINA
e-mail: ldlzj123@163.com

Abstract: In this paper, a variant of feasible descent SQP method is proposed to solve the nonlinear programming. In our algorithm, a feasible descent direction is obtained by solving only one modified QP subproblem. In order to avoid Maratos effect, a height-order correction direction is computed by an explicit formula. Under mild conditions without strict complementarity, the global and local superlinear convergence properties are obtained.

AMS Subject Classification: 65K05, 90C30
Key Words: inequality constrained optimization, SQP algorithm, global convergence, superlinear convergence, strict complementarity

1. Introduction

We consider the following inequality constrained nonlinear programming problem:

$$\begin{array}{ll} \min & f(x), \\ \text{s.t.} & g_j(x) \le 0, \quad j \in I = \{1, 2, \cdots, m\}, \end{array} \tag{1.1}$$

where $f : R^n \to R$, and $g_j : R^n \to R, (j \in I)$ are continuously differentiable.

Because of its superliner convergence rate, it is a topic of much active research [1]-[6]. In [6], Zhu presented an efficient SQP for problem (1.1). The basic feasible descent direction d is computed by solving the QP(x^k, H_k) prob-

Received: March 26, 2010
§Correspondence author

lem as follows:

$$\begin{array}{ll} \min & z + \frac{1}{2}d^T H_k d, \\ \text{s.t.} & \nabla f(x^k)^T d \le z, \\ & g_j(x^k) + \nabla g_j(x^k)^T d \le \eta_k z, \quad j \in I, \end{array} \tag{1.2}$$

where H_k is a symmetric positive definite matrix, and η_k is nonnegative scalar.

In this paper, another variant of feasible descent SQP method is proposed based on the subproblem proposed in [6]. The basic feasible descent direction is computed by solving the QP problem (1.2). But, unlike [6], in order to avoid Maratos effect, a height-order correction direction is computed by an explicit formula. Hence, under weaker assumptions without the strict complementarity, the algorithm is proved to possess global convergence, strong convergence and superlinear convergence.

2. Description of Algorithm

For the sake of simplicity, we denote

$$X = \{x \in R^n | g_j(x) \le 0, \ j \in I\}, \quad I(x) = \{j \in I | g_j(x) = 0\},$$

$$L_k = \{j \in I | g_j(x^k) + \nabla g_j(x^k)^T d^k = \eta_k z_k\}, \quad N_k := N(x^k) = (\nabla g_j(x^k), \ j \in I).$$

The following basic assumptions are assumed to be satisfied in this paper.

H 2.1. The feasible set is nonempty, i.e., $X \neq \emptyset$.

H 2.2. The functions f, $g_j(\ j \in I)$ are all continuously differentiable.

H 2.3. $\forall x \in X$, the set of vectors $\{\nabla g_j(x) : j \in I(x)\}$ is linearly independent.

For the QP(x^k, H_k)(1.2), if the matrix H_k is symmetric positive definite, (1.2) has a unique solution. And (d^k, z_k) is an optimal solution of (1.2) iff it is a KKT point of (1.2). Similar to the proof of Lemma 3.2 in [6], we can obtain the following result.

Lemma 2.1. *Suppose that Assumptions H2.1-H2.3 hold and (d^k, z_k) is an optimal solution of QP(x^k, H_k)(1.2). Then:*

1) $z_k \le -\dfrac{1}{2} d^{kT} H_k d^k \le 0$.

2) $z_k = 0 \Leftrightarrow d^k = 0 \Leftrightarrow k$ is a KKT point of (1.1).

3) If x^k is not a KKT point of the problem (1.1). Then $z_k = 0, \ d^k \neq 0$ and d^k is a feasible direction of descent for (1.1) at x^k.

To overcome the Maratoes effect, a suitable auxiliary direction must be adopted. In this paper, we use an explicit auxiliary formula to compute a auxiliary direction $\tilde{d^k}$ like in [4]:

$$\tilde{d^k} = -N_k(N_k^T N_k + D^k)^{-1}\big[\big(\|d^k\|^\tau + |\eta_k z_k|^\xi \|d^k\|\big)e + \tilde{g^k}\big], \tag{2.1}$$

where $e = (1, 1, \cdots 1)^T \in R^m$, $\tau \in (2, 3)$, $\xi \in (0, 1)$,

$$D_j^k = |g_j(x^k)|(|g_j(x^k) + \nabla g_j(x^k)^T d^k - \eta_k z_k| + |\eta_k z_k| + \|d^k\|), \tag{2.2}$$

$$\tilde{g^k}_j = g_j(x^k + d^k) - g_j(x^k) - \nabla g_j(x^k)^T d^k + \eta_k z_k. \tag{2.3}$$

From the assumptions H2.1-H2.3, it is easy to obtain the result that the formulas above are well defined. Now we state our algorithm as follows.

Algorithm A. *Step 0.* Given initial point $x^0 \in X$, a symmetric positive definite matrix $H_0 \in R^{n\times n}$. Choose parameters $\alpha \in (0, \frac{1}{2})$, γ, $\xi \in (0, 1)$, $\tau \in (2, 3)$, $\eta_0 = \rho > 0$. Set $k = 0$;

Step 1. Compute (d^k, z_k) by solving QP(x^k, H_k) subproblem (1.2). If $d^k = 0$ then STOP;

Step 2. Compute $\tilde{d^k}$ by (2.1). If $\|\tilde{d^k}\| > \|d^k\|$, then set $\tilde{d^k} = 0$.

Step 3. The line search. Compute t_k, the first number t in the sequence $\{1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \cdots\}$ satisfying:

$$f(x^k + td^k + t^2\tilde{d}^k) \le f(x^k) + \alpha t \nabla f(x^k)^T d^k, \tag{2.4}$$

$$g_j(x^k + td^k + t^2\tilde{d}^k) \le 0, \quad j \in I. \tag{2.5}$$

Step 4. Update. Generate H_{k+1}. Set $x^{k+1} = x^k + td^k + t^2\tilde{d}^k$, $\eta_{k+1} = \min\{\rho, \|d^k\|^\gamma\}$. Set $k := k + 1$. Go back to Step 1.

Under assumptions H 2.1-H 2.3, Algorithm A is well defined.

3. Global Convergence of Algorithm

In this section, we analyze the global convergence of the algorithm. The following general assumptions are true throughout this paper.

H 3.1. The sequences $\{x^k\}$ which is generated by the algorithm, is bounded.

H 3.2. There exist a, $b > 0$, such that $a\|d\|^2 \le d^T H_k d \le b\|d\|^2$, for all k and $d \in R^n$.

Lemma 3.1. *Suppose that Assumptions H 2.1-H 3.2 hold. Then:*

1) $\{\eta_k\}$ *and* $\{d^k,\ z_k,\ \nu_k,\ \mu^k\}$ *are bounded.*

2) There exists a constant $c_0 > 0$ *such that* $\|(N_k^T N_k + D^k)^{-1}\| \le c_0$.

3) The sequence $\{\tilde{d^k}\}$ *is bounded.*

Since there are only finitely many choices for set $L_k \subseteq I$ and due to the Lemma 3.1, we may assume that there exists an infinite index set K such that

$$x^k \to x^*,\ H_k \to H_*,\ d^k \to d^*,\ z_k \to z_*,\ \nu_k \to \nu_*,\ \mu^k \to \mu^*,\ \eta_k \to \eta_*,\ L_k \equiv L,\quad k \in K, \quad (3.1)$$

where L is a constant set.

Lemma 3.2. (see [6]) *Suppose that Assumptions H 2.1-H 3.2 hold,* K *is an infinite index set.*

1) If $\eta_k \to \eta_* > 0$, *then* $d^k \to 0$ $(k \in K)$.

2) Suppose $\{x^k\}$ *is a sequence generated by the algorithm,* $x^k \to x^*$, $d^k \to 0$, $(k \in K,\ k \to +\infty)$, *then* x^* *is a KKT point of (1.1).*

3) $(d^k, z_k) \to (0,0)$ $(k \in K,\ k \to +\infty)$.

Theorem 3.1. *Suppose that Assumptions H 2.1-H 3.2 hold. Then the proposed algorithm either stops at a KKT point for problem (1.1) in finite iterations, or generates an infinite sequence* $\{x^k\}$ *of points such that each accumulation* x^* *of* $\{x^k\}$ *is a KKT point for (1.1).*

Proof. The first statement is obvious, the only stopping point being in Step 1. Now, assume that the algorithm generates an infinite sequence $\{x^k\}$, and (3.1) holds. The cases $\eta_* = 0$, and $\eta_* > 0$ are considered separately.

If $\eta_* = 0$, in view of the definition of η_k, there exists an infinite index set $K' \subseteq K$, such that $d^{k-1} \to 0,\quad k \in K'$. From Step 3, it holds that

$$\|x^k - x^{k-1}\| \le t_k\|d^{k-1}\| + t_k^2\|\tilde{d^{k-1}}\| \le 2t_k\|d^{k-1}\| \to 0, \quad k \in K'.$$

So the fact that $x^k \to x^*$ implied that $x^{k-1} \to x^*$, $(k \in K')$. Then, we obtained

$$x^{k-1} \to x^*,\ d^{k-1} \to 0, \quad k \in K'.$$

Therefore, from the Lemma 3.2, it shows that x^* is a KKT point of (1.1).

If $\eta_* > 0$, it is clear that x^* is a KKT point of (1.1) from Lemma 3.2. □

4. Superlinear Convergence of Algorithm

In this section, we further discuss the strong convergence and superlinear convergence properties of the proposed algorithm under some mild conditions without the strict complementarity assumption. For this reason, we add the following additional hypothesis is necessary.

H 4.1. 1) The functions f and $g_j(x)$ $(j \in I)$ are all twice continuously differentiable.

2) The KKT pair (x^*, u^*) satisfies the strongly second-order sufficient conditions, i.e.

$$d^T \nabla^2_{xx} L(x^*, u^*) d > 0, \qquad \forall d \in \Omega \overset{\triangle}{=} \{d \in R^n | d \neq 0,\ \nabla g_{I_*^+}(x^*)^T d = 0\},$$

where

$$L(x,u) = f(x) + \sum_I u_j g_j(x), \quad I_*^+ = \{j \in I : u_j^* > 0\}.$$

Lemma 4.1. *For k large enough, it holds that:*

1) $\eta_k \to 0$, $\nu_k \to 1$, $\mu^k \to u^*$.

2) $(N_k^T N_k + D^k)^{-1} \to (N_*^T N_* + D^*)^{-1}$, $D^* = \mathrm{diag}\,(g_j(x^*)^2,\ j \in I)$.

3) $I_*^+ \subseteq L_k \subseteq I(x^*)$.

4) $z_k = O(\|d^k\|)$, $\|\tilde{d^k}\| = O(\|d^k\|^2) + O(|\eta_k z_k|) = o(\|d^k\|)$.

Proof. 1) The proof is similar to the Lemma 4.3 in [6].

2) In view of $(d^k, z_k) \to (0,0)(k \to +\infty)$ and the definition of D^k , we can get the matrix $(N_*^T N_* + D^*)$ is nonsingular and positive definite, furthermore

$$(N_k^T N_k + D^k)^{-1} \to (N_*^T N_* + D^*)^{-1}.$$

3) Taking into account

$$(x^k, \eta_k, d^k, z_k) \longrightarrow (x^*, 0, 0, 0), \quad k \to +\infty,$$

we obtain $L_k \subseteq I(x^*)$.

On the other hand, for $j \in I_*^+$, we get $\mu_j^k \to u_j^* > 0$ from 1), then $I_*^+ \subseteq L_k$ for k large enough.

4) In view of the first constraint of $\mathrm{QP}(x^k, H_k)$ (1.2), we have $\nabla f(x^k)^T d^k \leq z_k \leq 0$. Then $z_k = O(\|d^k\|)$. Taking into account the definition of $\tilde{d^k}$, it is easy to get $\|\tilde{d^k}\| = O(\|d^k\|^2) + O(|\eta_k z_k| \|d^k\|) + O(|\eta_k z_k|) = o(\|d^k\|)$. To ensure the unit step size can be accepted, we make another assumption as like H4.4 in [6].

Lemma 4.2. *For k large enough, the step $t_k \equiv 1$ is accepted by the Step*

3.

Proof. Firstly, we prove the inequalities of (2.5) holds for $t = 1$.

For $j \notin I(x^*)$, i.e., $g_j(x^*) < 0$. In view of $(x^k, d^k, z_k) \to (x^*, 0, 0)(k \to +\infty)$, we can conclude $g_j(x^k + d^k + \tilde{d^k}) \leq 0$ holds.

For $j \in I(x^*)$, $g_j(x^*) = 0$, then from (2.2), we have $D_j^k \to 0$ $(k \to \infty)$, furthermore,

$$|D_j^k| = o(|g_j(x^k) + \nabla g_j(x^k)^T d^k - \eta_k z_k|) + o(|\eta_k z_k|) + o(\|d^k\|). \qquad (4.1)$$

In view of Tarlor expansion, we have

$$\begin{aligned} g_j(x^k + d^k + \tilde{d^k}) &= g_j(x^k + d^k) + \nabla g_j(x^k + d^k)^T \tilde{d^k} + O(\|\tilde{d^k}\|^2) \\ &= g_j(x^k + d^k) + \nabla g_j(x^k)^T \tilde{d^k} + O(\|d^k\|^3) + O(|\eta_k z_k| \|d^k\|). \qquad (4.2) \end{aligned}$$

From (2.1)-(2.3) and (4.1) one has

$$\begin{aligned} \nabla g_j(x^k)^T \tilde{d^k} = &-\|d^k\|^\tau - |\eta_k z_k|^\xi \|d^k\| - g_j(x^k + d^k) + \nabla g_j(x^k)^T d^k - \eta_k z_k \\ &+ o(|g_j(x^k) + \nabla g_j(x^k)^T d^k - \eta_k z_k|) + o(\|d^k\|^3) + o(|\eta_k z_k| \|d^k\|). \qquad (4.3) \end{aligned}$$

Then, substituting (4.3) into (4.2) and considering the second constraint of (1.2), we get

$$\begin{aligned} g_j(x^k + d^k + \tilde{d^k}) = &-\|d^k\|^\tau - |\eta_k z_k|^\xi \|d^k\| + g_j(x^k) + \nabla g_j(x^k)^T d^k - \eta_k z_k \\ &+ o(|g_j(x^k) + \nabla g_j(x^k)^T d^k - \eta_k z_k|) + O(\|d^k\|^3) + O(|\eta_k z_k| \|d^k\|) \leq 0. \qquad (4.4) \end{aligned}$$

It implies that (2.5) holds for $t = 1$ and k large enough.

In addition, in view of (4.4), it is easy to get for $j \in L_k$

$$\begin{aligned} o(\|d^k\|^2) &= g_j(x^k + d^k + \tilde{d^k}) \\ &= g_j(x^k) + \nabla g_j(x^k)^T (d^k + \tilde{d^k}) + \frac{1}{2} d^{kT} \nabla g_j(x^k) d^k + o(\|d^k\|^2). \qquad (4.5) \end{aligned}$$

Secondly, we shows that a step of 1 brings a sufficient decrease on f, i.e., the inequalities of (2.4) holds for $t = 1$. The proof is similar the Lemma 4.4 in [6] and is omitted here. □

Imitating the analysis of Theorem 5.2 in [3], we can obtain the following result.

Theorem 4.1. *Under all above-mentioned assumptions, the algorithm is superlinearly convergent, i.e., the sequence* $\{x^k\}$ *generated by the algorithm satisfies*

$$\|x^{k+1} - x^*\| = o(\|x^k - x^*\|).$$

Acknowledgments

This work was supported in part by Scientific Research Fund of Hunan Institute of Humanities, Science and Technology of China (No. 2009QN07).

References

[1] J.F. Binnans, G. Launay, Sequential quadratic programming with penalization the displacement, *SIAM J. Optimization*, **54**, No. 4 (1995), 796-812.

[2] J.V. Burke, S.P. Han, A robust SQP method, *Math. Programming*, **43** (1989), 277-303.

[3] F. Facchinei, S. Lucidi, Quadraticly and superlinearly convergent for the solutionof inequality constained optimization problem, *JOTA*, **85**, No. 2 (1995), 265-289.

[4] J.B. Jian, A new feasible descent algorithm combining SQP with generalized projection for optimization problems without strict complementarity, *Appl. Math. Comput.*, **162** (2005), 1065-1081.

[5] C.T. Lawxence, A.L. Tits, Nonlinear equality constraints in feasible sequential quadratic programming, *Optim. Methods Softw.*, **6** (1996), 265-282.

[6] Z.B. Zhu, An efficient sequential quadratic programming algorithm for nonlinear programming, *J. Comput. Math.*, **175** (2005), 447-464.