Journal of Computational Mathematics, Vol.20, No.3, 2002, 301–312.

SEQUENTIAL SYSTEMS OF LINEAR EQUATIONS ALGORITHM FOR NONLINEAR OPTIMIZATION PROBLEMS – INEQUALITY CONSTRAINED PROBLEMS*1)

Zi-you Gao
(*School of Traffic and Transportation, Northern Jiaotong University, Beijing 100044, China*)

Tian-de Guo
(*Institute for Loo-Keng Hua Applied Mathematics and Information Sciences, Graduate School of Chinese Academy of Sciences, Beijing 100039, China*)

Guo-ping He   Fang Wu
(*Institute of Applied Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Abstract

In this paper, a new superlinearly convergent algorithm of sequential systems of linear equations (SSLE) for nonlinear optimization problems with inequality constraints is proposed. Since the new algorithm only needs to solve several systems of linear equations having a same coefficient matrix per iteration, the computation amount of the algorithm is much less than that of the existing SQP algorithms per iteration. Moreover, for the SQP type algorithms, there exist so-called inconsistent problems, i.e., quadratic programming subproblems of the SQP algorithms may not have a solution at some iterations, but this phenomenon will not occur with the SSLE algorithms because the related systems of linear equations always have solutions. Some numerical results are reported.

Key words: Optimization, Inequality constraints, Algorithms, Sequential systems of linear equations, Coefficient matrices, Superlinear convergence.

1. Introduction

In this paper, we consider the following nonlinear programming problem with inequality constraints:

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

where $x = (x_1, \cdots, x_n)^T \in E^n$, the index set $I = \{1, 2, \cdots, m\}$ and $f: E^n \to E$ and $g_j: E^n \to E$ $(j = 1, \cdots, m)$ are all real-valued functions.

Since the algorithms of Sequential Quadratic Programming (i.e. SQP) generally have good superlinear convergence properties, they are currently considered to be one of the most effective approaches for solving nonlinear programming problems with nonlinear constraints, and have been widely studied by many authors. (See, e.g. [5-8]). However, most of the SQP algorithms have two serious shortcomings: (1) In order to obtain a search direction, one must solve one or more quadratic programming subproblems per iteration, general speaking, the computation amount of this type of algorithms is very large. In addition, it is difficult to use some good

* Received May 23, 1999; Final revised March 28, 2001.
1)This research was supported by the National Natural Science Foundation of China(19571001, 19971002, 79970014) and Cross-century Excellent Personnel Award and Teaching and Research Award Program for Outstanding Young Teachers in High Education, Ministry of Education, China.

sparse and symmetric properties in solving quadratic programming subproblems, this may restrict the application of the SQP type algorithms, especially for large scale problems; (2) The SQP algorithms require that the related quadratic programming subproblems must be solvable per iteration, however, it is difficult to be satisfied in general. So it is desirable to design some algorithms which can avoid these shortcomings for nonlinear optimization problems.

In Ref.[3], E.R. Panier, A,L.Tits and J.N. Herskovits gave a feasible algorithm which is two-step superlinear convergent. They try to replace quadratic programming problems by systems of linear equations for overcoming the difficulties encountered in the SQP methods. The algorithm in Ref.[3] needs to solve two linear systems and a quadratic subproblem at each iteration, and the initial point of the algorithm must be an interior point. Particularly, in order to obtain the global convergence, Ref.[3] needs two strong assumptions (i.e. an interior point is assumed, and the number of stationary points is finite.). In addition, the algorithm of Ref.[3] can not be used to deal with problems having equality constraints.

In order to overcome these shortcomings just mentioned for the SQP algorithms for solving constrained optimization problems, through improving the algorithm in Ref.[3], a new superlinearly convergent algorithm of sequential systems of linear equations (SSLE) for inequality constrained optimization problems is proposed in this paper. Comparing with the SQP algorithms recently suggested, the new algorithm has three main advantages: (1) The new algorithm is completely QP-free, more precisely, the algorithm only needs to solve four systems of linear equations having a same coefficient matrix per iteration, the computation amount of the algorithm is much less than that of the existing SQP algorithms per iteration; (2) The iterating points generated by the algorithm are feasible; (3) The rate of convergence is one-step superlinear under much milder assumptions than that in Ref.[3].

The proposed method is based on the following observation. Let $\{d_k^0\}$ be a sequence generated by the following linear system in (d, λ):

$$H_k d_k^0 + \nabla f(x_k) + \sum_{j=1}^{m} \lambda_{k,j}^0 \nabla g_j(x_k) = 0, \tag{1.1}$$

$$\mu_{k,j} \nabla g_j(x_k)^T d_k^0 + \lambda_{k,j}^0 g_j(x_k) = 0, \qquad (j \in I), \tag{1.2}$$

where H_k is an estimate of the Hessian of $L(x, \lambda) = f(x) + \sum_{j=1}^m \lambda_j g_j(x)$, x_k the current estimate of a solution x^*, $x_k + d_k^0$ the next estimate, μ_k the current estimate of the Kuhn-Tucker multiplier vector associated with x^*, and λ_k^0 the next estimate of this vector. Locally, the system (1.1)-(1.2) is a higher order perturbation of the following quadratic programming (SQP):

$$\begin{aligned} min \quad & \frac{1}{2} d^T H_k d + \nabla f(x_k)^T d \\ s.t. \quad & g_j(x_k) + \nabla g_j(x_k)^T d \le 0, \qquad j = 1, \cdots, m, \end{aligned}$$

according to the theorem 4.6 of Ref.[3], we know that the sequence defined by $x_{k+1} = x_k + d_k^0$ have superlinear convergence. Thus, d_k^0 is obtained as first direction. However, d_k^0 is not entirely suitable as a search direction. Indeed, although d_k^0 is a descent direction, it may be zero at some iterates which are not K-T point of (IP). This effect can be avoided if one substitutes in each right-hand side of (1.2) a negative number v_j, $(j \in I)$, but v_j must tend to zero faster than d_k^0 in order for the convergence properties of $\{x_k\}$ to be preserved, we let v_j be $(\lambda_{k,j}^0)$ or $(-\lambda_{k,j}^0) g_j(x_k) ||d_k^0||^2$. Thus, after d_k^0 is obtained, a new direction d_k^1 would be computed by solving the linear system

$$\begin{aligned} & H_k d_k^1 + \nabla f(x_k) + \sum_{j=1}^{m} \lambda_{k,j}^1 \nabla g_j(x_k) = 0, \\ & \mu_{k,j} \nabla g_j(x_k)^T d_k^1 + \lambda_{k,j}^1 g_j(x_k) = \mu_{k,j} v_{k,j}, \qquad (j \in I), \end{aligned} \tag{1.3}$$