Journal of Computational Mathematics, Vol.19, No.4, 2001, 337–346.

GLOBAL CONVERGENCE AND IMPLEMENTATION OF NGTN METHOD FOR SOLVING LARGE-SCALE SPARSE NONLINEAR PROGRAMMING PROBLEMS*[1)]

Qin Ni
(*Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*)

Abstract

An NGTN method was proposed for solving large-scale sparse nonlinear programming (NLP) problems. This is a hybrid method of a truncated Newton direction and a modified negative gradient direction, which is suitable for handling sparse data structure and possesses Q-quadratic convergence rate. The global convergence of this new method is proved, the convergence rate is further analysed, and the detailed implementation is discussed in this paper. Some numerical tests for solving truss optimization and large sparse problems are reported. The theoretical and numerical results show that the new method is efficient for solving large-scale sparse NLP problems.

Key words: Nonlinear programming, Large-scale problem, Sparse.

1. Introduction

Consider the following NLP problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g_j(x) \geq 0, \quad j \in J = \{1, \ldots, m\}. \end{array} \tag{1.1}$$

where the function $f : R^n \to R^1$ and $g_j : R^n \to R^1,\ j \in J$ are twice continuously differentiable. In particular, we discuss the case, where the number of variables and the number of constraints in (1.1) are large and second derivatives in (1.1) are sparse.

There are some methods which can solve large-scale problems, e.g. Lancelot in [2] and TDSQPLM in [9]. But they can not take advantage of sparse structure of the problem. A new efficient method which is called NGTN method is studied in [11] for solving large-scale sparse NLP problems. In this method, a new nonlinear system which is equivalent to Kuhn-Tucker conditions of the problem is developed. NCP function is used in the nonlinear system such that the nonnegativity of some variables is avoided. The truncated Newton method is used to solve the nonlinear system. In order to guarantee the global convergence, a robust loss function is chosen as a merit function and a modified negative gradient direction is used to descrease the merit function. This NGTN method is easy to carry out, possesses Q-quadratic convergence rate, and is suitable for solving large-scale sparse NLP problems.

In this paper, the global convergence of NGTN method is proved, Q-quadratical convergence rate is further analysed, and the detailed implementation is discussed. In addition, NGTN algorithm is used for solving the sparse truss problems, the dimensions of which range from 183 to 827, and large problem with 70000 variables and 40000 constraints. The theoretical and numerical results show that NGTN method is efficient for solving large-scale sparse NLP

* Received April 9, 1998; Final revised December 2, 1998.
[1)] This research was supported by Nationa Natural Science Foundation of China, LSEC of CAS in Beijing and Natural Science Foundation.

problems. This paper is organized as follows. In Section 2 we give the construction of a NGTN algorithm for solving the large-scale problem (1.1). We discuss the global convergence of NGTN in Section 3. The detailed implementation and numerical results of NGTN are given in Section 4.

2. NGTN Algorithm

2.1. Notations

In order to describe the NGTN algorithm, let

$$L(x,u) = f(x) - \sum_{j=1}^{m} u_j g_j(x) \tag{2.1}$$

be the Lagrangian function of problem (1.1), let $\nabla_x L(x,u)$ and $\nabla^2_{xx} L(x,u)$ denote the gradient and Hessian of $L(x,u)$ at x, respectively. With this notation, a pair (x^*, u^*) is called a Kuhn-Tucker pair of (1.1) if (x^*, u^*) satisfies the following Kuhn-Tucker conditions:

$$\nabla_x L(x^*, u^*) = 0, \ g_j(x^*) - t_j^* = 0, \ j \in J,$$
$$u_j^* t_j^* = 0, \ u_j^* \geq 0, \ t_j^* \geq 0, \ j \in J.$$

The Kuhn-Tucker conditions are equivalent to the system

$$0 = \quad q(z) = \begin{bmatrix} \nabla_x L(x,u) \\ --------- \\ t_j - g_j(x) \\ --------- \\ \phi_j(u,t) \end{bmatrix}_{j \in J} \tag{2.2}$$

where $q : R^{n+2m} \to R^{n+2m}, z = (x,u,t)$ and

$$\phi_j(u,t) = \sqrt{u_j^2 + t_j^2} - (u_j + t_j) \tag{2.3}$$

is a NCP-function (see [3]).

The Jacobian matrix of $q(z)$ is

$$Q(z) = \begin{pmatrix} \nabla^2_{xx} L(x,u) & -A^T(x) & 0 \\ -A(x) & 0 & I_m \\ 0 & \Phi_1(u,t) & \Phi_2(u,t) \end{pmatrix}, \tag{2.4}$$

where

$$\begin{aligned} A(x) &= (\nabla g_1(x), \ldots, \nabla g_m x))^T \in R^{m \times n}, \\ \Phi_1(u,t) &= \mathrm{diag}(\frac{\partial \phi_1(u,t)}{\partial u_1}, \ldots, \frac{\partial \phi_m(u,t)}{\partial u_m}), \\ \Phi_2(u,t) &= \mathrm{diag}(\frac{\partial \phi_1(u,t)}{\partial t_1}, \ldots, \frac{\partial \phi_m(u,t)}{\partial t_m}). \end{aligned}$$

$(\Delta x_k, \Delta \tilde{u}_k)$ is called a truncated solution, if the following equation

$$\begin{pmatrix} H_k & -\tilde{A}_k^T \\ -\tilde{A}_k & -D_k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \tilde{u} \end{pmatrix} = \begin{pmatrix} b_{k1} \\ b_{k2} \end{pmatrix} \tag{2.5}$$

is solved such that the inequality

$$\| \begin{pmatrix} H_k & -\tilde{A}_k^T \\ -\tilde{A}_k & -D_k \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \tilde{u} \end{pmatrix} - \begin{pmatrix} b_{k1} \\ b_{k2} \end{pmatrix} \| \leq \eta_k \tag{2.6}$$

holds for some $\eta_k > 0$. Here

$$\tilde{A}_k^T = (\nabla g_j(x_k))_{j \in J_k} \in R^{n \times m_k}, J_k = \{j \in J : \ |\eta_j^{(k)}| \geq \epsilon_1 \ \text{ or } \ |\xi_j^{(k)}| \leq \epsilon_2\} \tag{2.7}$$