Journal of Computational Mathematics, Vol.20, No.3, 2002, 225–244.

AN INTERIOR TRUST REGION ALGORITHM FOR NONLINEAR MINIMIZATION WITH LINEAR CONSTRAINTS*[1)]

Jian-guo Liu

(*Department of Mathematics, University of North Texas, Denton, TX 76203, USA*)

Abstract

An interior trust-region-based algorithm for linearly constrained minimization problems is proposed and analyzed. This algorithm is similar to trust region algorithms for unconstrained minimization: a trust region subproblem on a subspace is solved in each iteration. We establish that the proposed algorithm has convergence properties analogous to those of the trust region algorithms for unconstrained minimization. Namely, every limit point of the generated sequence satisfies the Krush-Kuhn-Tucker (KKT) conditions and at least one limit point satisfies second order necessary optimality conditions. In addition, if one limit point is a strong local minimizer and the Hessian is Lipschitz continuous in a neighborhood of that point, then the generated sequence converges globally to that point in the rate of at least 2-step quadratic. We are mainly concerned with the theoretical properties of the algorithm in this paper. Implementation issues and adaptation to large-scale problems will be addressed in a future report.

Key words: Nonlinear programming, Linear constraints, Trust region algorithms, Newton methods, Interior algorithms, Quadratic convergence.

1. Introduction

Consider the following linearly constrained minimization problem:

$$\min_{x\in\Re^n} f(x) : \Re^n \Rightarrow \Re \tag{1.1}$$

$$\text{subject to } Ax = b \text{ and } x \geq 0,$$

where f is assumed to be twice continuously differentiable, $A \in \Re^{m\times n}$, and $b \in \Re^m$. We are interested in locating a local minimizer and call such a minimizer a solution to (1.1).

We propose an interior trust region based algorithm. Starting from a strictly feasible point x^0 (or, interior point, i.e., $Ax^0 = b$ and $x^0 > 0$), a sequence $\{x^k\}$ is generated and every x^k remains interior. In each iteration, a trust region subproblem is solved and the iterate x^k is updated. A projected gradient is used as a watch dog to guarantee global convergence. Under certain assumptions, we establish that the algorithm has convergence properties analogous to those of the trust region algorithms for unconstrained minimization (see, e.g., [28]). Namely,

* Received March 23, 1999; Final revised September 26, 2001.

1) Research partially supported by the Faculty Research Grant RIG-35547 and ROG-34628 of the University of North Texas, and in part by the Cornell Theory Center which receives major funding from the National Science Foundation and IBM Corporation, with additional support from New York State and members of its Corporate Research Institute. This work was partially done while the author was visiting Cornell University and the University of California, San Diego.

every limit point of the generated sequence satisfies the Krush-Kuhn-Tucker (KKT) conditions and at least one limit point satisfies second order necessary optimality conditions. In addition, if one limit point is a strong local minimizer and the Hessian of f is Lipschitz continuous in a neighborhood of that point, then the sequence converges globally to that point and the rate of convergence is at least 2-step quadratic.

Trust region algorithms for unconstrained minimization have been studied by many authors (see, e.g., [14], [21], [23], [27], and [28]). A trust region algorithm can be briefly described as follows.

Let x be the current approximation to a solution of $\min_x f(x)$ and let $\delta > 0$ be the current trust region radius. A solution to the model trust region subproblem

$$\min_{\Delta x} \{q(\Delta x) := f + \nabla f^T \Delta x + \frac{1}{2}\Delta x^T \nabla^2 f \Delta x \; : \; \|\Delta x\| \leq \delta\} \tag{1.2}$$

is computed. If the actual reduction on the objective function, $f(x) - f(x+\Delta x)$, is satisfactory comparing with $f(x) - q(\Delta x)$, the reduction predicted by the quadratic model q in (1.2), $x + \Delta x$ is taken as the next approximation, and the trust region radius δ is updated for the next iteration. Otherwise the trust region radius is reduced and a new trust region subproblem is solved. A reduction is satisfactory if

$$\frac{f(x) - f(x+\Delta x)}{f(x) - q(\Delta x)} \geq \eta,$$

where $\eta \in (0, 1)$ is a given constant.

Excellent global convergence properties of the trust region algorithms for unconstrained minimization have been established. In [28] for example, it is shown that every limit point of the generated sequence satisfies first order conditions and at least one limit point satisfies second order necessary optimality conditions. In addition, if there is one limit point which is a strong local minimizer, then the whole sequence converges globally to this point and the rate of convergence is quadratic. In [27], it is proved that if a reasonable strategy for increasing the trust region radius is imposed, then every limit point satisfies second order necessary conditions.

Trust region algorithms have also been applied to minimization problems with simple bounds (see, e.g., [6] and [10]), to equality constrained minimization (see, e.g., [2], [4], [24], and [29]), and to linearly constrained minimization (see, e.g., [5], [11], and [15]). Similar convergence results are also obtained.

Among the algorithms for linearly constrained minimization, many are based on the active set method, such as [3], [12], [13], [16], [18], [19], [22], [25], [26], and [30]. But some are of interior point type, e.g., [11] and [15].

Interior algorithms have been extensively studied and applied to many optimization problems, including linear programming, linear complementarity problems, and quadratic programming. A main feature of the interior algorithms is that the number of iterations is not very sensitive to problem dimension. For references, see, e.g., [20], [31], and [32].

This paper is organized as follows. In Section 2, we first give the optimality conditions for (1.1) and then motivate and define our algorithm. Convergence properties are established in Section 3 and 4. Finally, we have some concluding remarks in Section 5.

We use $\nabla f(x)$ to denote the gradient of $f(x)$ and $\nabla^2 f(x)$ to denote the Hessian. Given x^* and x^k, we write $f^* = f(x^*)$, $f^k = f(x^k)$, and use similar notations for ∇f and $\nabla^2 f$. We use superscripts to denote the iteration counts and use subscripts to indicate the indices of vector