Journal of Computational Mathematics, Vol.19, No.2, 2001, 113–124.

ON MAXIMA OF DUAL FUNCTION OF THE CDT SUBPROBLEM*[1)]

Xiong-da Chen
(*Reserach Development Center of Parallel Software, Institute of Software, Beijing 100080, China*)

Ya-xiang Yuan
(*State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Abstract

In this paper, we show the geometry meaning of the maxima of the CDT subproblem's dual function. We also studied the continuity of the global solution of the trust region subproblem. Based on an approximation model, we prove that the global solution of the CDT subproblem is given with the Hessian of Lagrangian positive semi-definite by some specially-located dual maxima and by restricting the location region of the multipliers which corresponding a global solution in other cases.

Key words: Trust region subproblem, Global minimizer, Approximation.

1. Introduction

Consider the following the CDT problem P

$$\min_{d\in\mathcal{R}^n} \Phi(d) = \frac{1}{2}d^TBd + g^Td \tag{1.1}$$

subject to

$$\|d\| \le \Delta, \tag{1.2}$$

$$\|A^Td + c\| \le \xi, \tag{1.3}$$

where $g \in \mathcal{R}^n$, $B \in \mathcal{R}^{n\times n}$, $A \in \mathcal{R}^{n\times m}$, $c \in \mathcal{R}^m$, $\Delta > 0$, $\xi \ge 0$, B is a symmetric matrix not necessary positive semi-definite, and throughout this paper, the norm $\|\cdot\|$ denotes the Euclidean norm. For the convenient of our following discussion, let $\mathcal{F}$ be the feasible region of the CDT subproblem,

$$\mathcal{F}_0 = \{d \mid \|A^Td + c\| \le \xi\}, \tag{1.4}$$

and

$$\mathcal{F}_1 = \{d \mid \|d\| \le \Delta\}. \tag{1.5}$$

Problem (1.1)–(1.3) arises in some trust region algorithms for equality constrained optimization aiming to conquer the inconsistency between the trust region and the linearized constraints of original problem in every iteration. Called the CDT subproblem, it was first proposed by Celis, Dennis and Tapia (1985), and later it was applied in algorithms for equality constrained optimization to achieve certain property of global convergence, for example, see Powell and Yuan (1991).

The CDT subproblem is often required to compute the global solution or to satisfy some kind of sufficient descent property in some algorithm for the equality constrained optimization, for instance, see also Powell and Yuan (1991). If B is positive semi-definite, problem (1.1)–(1.3) is a convex optimization. Several authors studied its properties and then proposed algorithms to find its global minimizer respectively, for example, see Yuan (1991) and Zhang (1992). If B

* Received April 7, 1998.
[1)] Research partially supported by Chinese NSF grants 19525101, 19731010 and State key project 96-221-04-02-02.

is not positive semi-definite, unlike the following trust region subproblem P_1 with single ball constraint which is of form

$$\min_{d\in\mathcal{R}^n} \Phi(d) = \frac{1}{2}d^TBd + g^Td \tag{1.6}$$

subject to

$$\|d\| \leq \Delta, \tag{1.7}$$

the CDT subproblem has still no satisfactory algorithm to find a global solution.

Now we introduce some notations about dual variables and dual function. Using the notations of Yuan (1991), we define the Hessian of Lagrangian

$$H(\lambda,\mu) = B + \lambda I + \mu AA^T, \tag{1.8}$$

where $\lambda \geq 0$, $\mu \geq 0$ are the Lagrangian multipliers. If $H(\lambda,\mu)$ is nonsingular, we define the vector

$$d(\lambda,\mu) = -H(\lambda,\mu)^{-1}(g + \mu Ac) \tag{1.9}$$

and the Lagrangian dual function

$$\Psi(\lambda,\mu) = \Phi(d) + \frac{\lambda}{2}(\|d\|^2 - \Delta^2) + \frac{\mu}{2}(\|A^Td + c\|^2 - \xi^2), \tag{1.10}$$

where $d = d(\lambda,\mu)$ is given above. Thus, the Lagrangian multipliers are also the dual variables.

It is well known that, without assuming the positive semi-definiteness of B, $d(\lambda,\mu)$ given by (1.9) is the global solution of (1.1)–(1.3) if $H(\lambda,\mu)$ is positive definite at a maxima of the dual function(for example, see Yuan (1990)). But if the maxima locates on the boundary of the positive semi-definite region

$$\Omega_0 = \{(\lambda,\mu) \in \mathcal{R}^2_+ \mid H(\lambda,\mu) \text{ is positive semidefinite }\}, \tag{1.11}$$

where, $\mathcal{R}^2_+ = \{\lambda \geq 0, \mu \geq 0\}$, what does the dual variables correspond? And what property does it possess? In this paper, we discuss the geometry meaning of the maxima through the insight of continuity of the global solution of the single ball constrained subproblem.

Without the assumption of the positive semi-definiteness of B, we will give the more detailed properties of single-ball-constrained trust region subproblem and show the geometry meaning of the dual maxima of the CDT problem on the region Ω_0. And we extend the result of location of multipliers which corresponding the global solution.

The paper is organized as follows: in section 2, we present some properties of trust region subproblem (1.6)–(1.7). In section 3 and section 4, we construct an approximation of the feasible region of the CDT problem, which forms a new trust region subproblem with a parameter w, and then we discuss the relations between the CDT problem and the new trust region subproblem. In section 5, we illustrate the geometry meaning of a certain parameter of the new trust region subproblem, which we call it a jump. We also strengthen the result in Chen and Yuan (1998) by further studying the dual maxima of the CDT problem in last section.

2. Properties of Trust Region Subproblem

In this section, we study the global solution of the trust region subproblem, which has the form of (1.6)–(1.7) where B is a symmetric matrix. We first introduce a theorem which characterizes the global solution of problem (1.6)–(1.7) which is given independently by Gay (1981) and Sorensen (1982), see also Moré and Sorensen(1983).

Theorem 2.1. *A feasible point $d^* \in \mathcal{R}^n$ is the global solution to problem (1.6)–(1.7), if and only if there exists a $\lambda^* \geq 0$ such that*

$$(B + \lambda^* I)d^* = -g \tag{2.1}$$

and

$$\lambda^*(\|d^*\| - \Delta) = 0, \tag{2.2}$$

where $B + \lambda^* I$ is positive semidefinite.