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.

We now study the so-called hard case of problem P_1, i.e., there is no $\lambda \geq 0$ such that $B + \lambda I$ is positive definite and

$$\|(B + \lambda I)^{-1} g\| = \Delta, \tag{2.3}$$

for more details of known results, see Moré and Sorensen(1983). In the hard case, to find a global minimizer of problem P_1 is equivalent to find a unit vector $z \in \mathcal{N}(B - \rho_1 I)$ and determine a scalar τ such that

$$(B - \rho_1 I)(p + \tau z) = -g, \tag{2.4}$$

and

$$\|p + \tau z\| = \Delta, \tag{2.5}$$

where p solves

$$(B - \rho_1 I)p = -g. \tag{2.6}$$

Here, $\rho_1 \leq 0$ is the smallest eigenvalue of B and $\mathcal{N}(B)$ denotes the null space of B.

In the following, we denote $dim(\cdot)$ the dimension of a space. Suppose that there is no $\lambda \geq 0$ such that (2.3) holds with $B + \lambda I$ positive definite, i.e., the hard case holds, we can divide it into following cases :

- case 1). if $\|(B - \rho_1 I)^+ g\| = \Delta$, then the global solution is $p = -(B - \rho_1 I)^+ g$, the least square solution of equation (2.6). Here B^+ means the Moore-Penrose general inverse of B. So, in fact, we need not solve the equations (2.4)–(2.5), where $\tau = 0$. In further studies, we will see that this is different to the other hard case.
- case 2). if $dim(\mathcal{N}(B - \rho_1 I)) = 1$, then for a given vector $z \in \mathcal{N}(B - \rho_1 I)$, $\|z\| = 1$, the equation (2.5) of τ will have two solutions. The the global solutions are of form $-(B - \rho_1 I)^+ g \pm \tau z$.
- case 3). if $dim(\mathcal{N}(B - \rho_1 I)) \geq 2$, the global solutions have the form $-(B - \rho_1 I)^+ g + \tau z$ with $z \in \mathcal{N}(B - \rho_1 I)$ and $\|z\| = 1$. Since $z \in \mathcal{N}(B - \rho_1 I)$ and $-(B - \rho_1 I)^+ g \in \mathcal{R}(B - \rho_1 I)$, all the global solutions share the same value of τ. Here, $\mathcal{R}(\cdot)$ denotes the range space of a matrix. The set of global solutions actually is a lower dimensional sphere in $\mathcal{R}^n$.

Since the global solution of problem P_1 is a single point when the corresponding Hessian is positive definite or $\|(B - \rho_1 I)^+ g\| = \Delta$ in the so-called hard case, the set of global solutions is a disconnected set when the second case holds. Hence we have

Theorem 2.2. *The set of global solutions of problem P_1 is disconnect if and only if*
i) *there is no $\lambda \geq 0$ with $B + \lambda I$ positive definite such that (2.3) holds, and*
$\|(B - \rho_1 I)^+ g\| < \Delta$;
ii) *ρ_1 is the single-tuple eigenvalue of B.*

In the following, we discuss the continuity of the global solution of problem P_1. First, we define the distance of a point x and a set S. Define that

$$dist(x, S) = \inf_{y \in S} dist(x, y), \tag{2.7}$$

where $dist(x, y) = \|x - y\|$. We have the fact that the set of global solutions of problem P_1 continuously depends on B and g, which can be stated as the following theorem.

Theorem 2.3. *Suppose that S is the set of global solutions of problem P_1, $\tilde{S}$ is the set of global solutions of following perturbed problem*

$$\min_{d \in \mathcal{R}^n} \{\tilde{\Phi}(d) \mid d \in \mathcal{F}_1\}, \tag{2.8}$$

where $\tilde{\Phi}(d) = \frac{1}{2} d^T \tilde{B} d + \tilde{g}^T d$. Then, $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$, such that for all $\tilde{B} \in \mathcal{R}^{n \times n}$ and $\tilde{g} \in \mathcal{R}^n$, if $max(\|\tilde{B} - B\|, \|\tilde{g} - g\|) < \delta$, we have

$$dist(\tilde{d}, S) \leq \varepsilon, \tag{2.9}$$

for all $\tilde{d} \in \tilde{S}$.

Proof. Let v^* be the global optimal value of problem P_1, d^* is a global minimizer, hence $\Phi(d^*) = v^*$. From the continuity of problem P_1, S is a closed set, and so is $\tilde{S}$. Then, for $\eta > 0$ sufficiently small,

$$S_\eta = \{d \mid \Phi(d) < v^* + \eta,\ \|d\| \le \Delta\} \tag{2.10}$$

is an open neighborhood of S (set the feasible region of problem P_1 as the whole space). It is easy to see that $d^* \in S_\eta$. Moreover, $\forall\, \varepsilon > 0$, $\exists\, \eta > 0$, such that

$$dist(d, S) \le \varepsilon,\ \forall d \in S_\eta. \tag{2.11}$$

Otherwise, we have, $\exists\, \varepsilon > 0, \forall\, \eta > 0$, there exists a $d_\eta \in S_\eta$ such that $dist(d_\eta, S) > \varepsilon$. Because S and S_η are bounded closed sets, let $\eta \to 0$ and d_0 be a cluster point of $\{d_\eta\}$, then $dist(d_0, S) > 0$ and $\Phi(d_0) = v^*$, this contradicts to the definition of S. There exists a $\delta > 0$, $\delta \le \frac{\eta}{3}(\Delta^2 + 2\Delta)^{-1}$ such that, if $max(\|\tilde{B} - B\|, \|\tilde{g} - g\|) < \delta$, then for all $d \in \mathcal{F}_1$,

$$\begin{aligned} \left|\tilde{\Phi}(d) - \Phi(d)\right| &= \left|\tfrac{1}{2} d^T(\tilde{B} - B)d + (\tilde{g} - g)^T d\right| \\ &\le \tfrac{1}{2}\delta\Delta^2 + \delta\Delta \\ &\le \frac{\eta}{3}. \end{aligned} \tag{2.12}$$

For all $d \in \mathcal{F}_1 \setminus S_\eta$, since $\Phi(d) \ge v^* + \eta$,

$$\tilde{\Phi}(d) \ge v^* + \eta - \frac{\eta}{3} = v^* + \frac{2}{3}\eta \tag{2.13}$$

implied by (2.12). Denote $\tilde{v}^*$ the global optimal value of problem (2.8), for the same reason, we have

$$\tilde{v}^* \le \tilde{\Phi}(d^*) \le \Phi(d^*) + \frac{\eta}{3} = v^* + \frac{\eta}{3}. \tag{2.14}$$

Then, $\tilde{S} \cap (\mathcal{F}_1 \setminus S_\eta) = \emptyset$ implied by (2.13) and (2.14). Thus the optimal point of (2.8) $\tilde{d} \in S_\eta$, and $dist(\tilde{d}, S) \le \varepsilon$.

3. An Approximation of the CDT Feasible Region

In this section, we construct a kind of approximation of feasible region $\mathcal{F}$ of problem P_1. There are some other approaches to construct an approximation of the region $\mathcal{F}$, for more detail, see El-Alem and Tapia (1995) and Fu, Luo and Ye (1996). We now investigate the global solution of the "scaled" problem P_w which has the following form:

$$\min_{d \in \mathcal{R}^n} \Phi(d) = \frac{1}{2} d^T B d + g^T d \tag{3.1}$$

subject to

$$w(\|d\|^2 - \Delta^2) + (1 - w)(\|A^T d + c\|^2 - \xi^2) \le 0, \tag{3.2}$$

where, $w \in [0, 1]$. The feasible region of problem P_w is $\mathcal{F}_0, \mathcal{F}_1$ when $w = 0, 1$ respectively. It is also easy to show that for any $w \in [0, 1]$, $\mathcal{F}_w$, the feasible region of P_w includes $\mathcal{F}$, the feasible region of the CDT subproblem, and is included in $\mathcal{F}_1 \cup \mathcal{F}_0$, i.e.,

$$\mathcal{F}_1 \cap \mathcal{F}_0 = \mathcal{F} \subseteq \mathcal{F}_w \subseteq \mathcal{F}_1 \cup \mathcal{F}_0. \tag{3.3}$$

Direct calculations show that problem (3.1)–(3.2) is equivalent to problem

$$\min_{d \in \mathcal{R}^n} \Phi(d) = \frac{1}{2} d^T B d + g^T d \tag{3.4}$$

subject to

$$\|\tilde{D}^T d + \tilde{c}\| \le \tilde{\xi}, \tag{3.5}$$

where,

$$\begin{cases} \tilde{D} = (wI + (1 - w)AA^T)^{\frac{1}{2}}, \\ \tilde{c} = (1 - w)\tilde{D}^{-1}Ac, \\ \tilde{\xi} = (\tilde{c}^T\tilde{c} + w\Delta^2 - (1 - w)(\xi^2 - c^T c))^{\frac{1}{2}}. \end{cases} \tag{3.6}$$

For $w \in [0,1]$, denote $S(w)$ the set of global solutions of the scaled problem P_w with parameter w, where $S(0)$ may be an empty set. If AA^T is nonsingular, or if $w \in (0,1]$, $\tilde{D}$ is positive definite and the problem P_w is equivalent to the transformed problem

$$\min_{\tilde{d}\in\mathcal{R}^n} \frac{1}{2}\tilde{d}^T\tilde{B}\tilde{d} + \tilde{g}^T\tilde{d} \tag{3.7}$$

subject to

$$\|\tilde{d}\| \leq 1, \tag{3.8}$$

where

$$\tilde{B} = \tilde{\xi}^2\tilde{D}^{-1}B\tilde{D}^{-T} \quad \text{and} \quad \tilde{g} = \tilde{\xi}\tilde{D}^{-1}(g - B\tilde{D}^{-T}c). \tag{3.9}$$

Given a feasible point or a solution $\tilde{d}$ of problem (3.7)–(3.8), the corresponding feasible point or solution of problem (3.4)–(3.5) is

$$d = \tilde{D}^{-T}(\tilde{\xi}\tilde{d} - \tilde{c}). \tag{3.10}$$

Since (3.10) is a one-to-one continuous mapping $\tilde{d} \to d$ defined on the bounded closed set $\{\|\tilde{d}\| \leq 1\}$, the following lemma follows from Theorem 2.3:

Lemma 3.1. *Let $S(w)$ be the set of global solutions of P_w, for all $\varepsilon > 0$, there exists a $\delta > 0$, such that*

$$dist(S(w_1), S(w_2)) < \varepsilon, \tag{3.11}$$

for any $w_1, w_2 \in (0,1]$, $|w_1 - w_2| < \delta$.

From the above analysis, if AA^T is nonsingular, the interval $(0,1]$ can be extended to $[0,1]$. But if AA^T is singular, P_0 can not be transformed to problem of form (3.7)–(3.8), moreover, it may even have no global solution, i.e., its global optimal objective value can be $-\infty$. On the other hand, for any fixed small $w > 0$, problem P_w has a finite optimal objective value. However, under this circumstance, we prove a property of problem P_0, i.e.,

$$\min\{\Phi(d) \mid d \in \mathcal{F}_0\}, \tag{3.12}$$

which, roughly speaking, is weaker than the result of lemma 3.1.

Lemma 3.2. *If $S(0) \cap \mathcal{F}_1 = \emptyset$, then for all $w > 0$ sufficiently small, $S(w) \cap \mathcal{F}_1 = \emptyset$.*

Proof. $S(0) \cap \mathcal{F}_1 = \emptyset$ means that for all global optimal solution d_0^* of problem P_0, $d_0^* \notin \mathcal{F}_1$. We prove the result by contradiction. Suppose the opposite case holds, then there exists a sequence $w \to 0+$, such that $S(w) \cap \mathcal{F}_1 \neq \emptyset$, which means there exists $d_w \in \mathcal{F}_1$ is a global solution of problem P_w, i.e., $d_w \in S(w)$.

Let d_0 be a cluster point of $\{d_w\}$, then $d_0 \in \mathcal{F}_1$. Without loss of generality, we assume that $d_w \to d_0$. Since d_w is a feasible point for problem P_w, thus (3.2) holds. Let $w \to 0+$, recalling that $d_w \to d_0$, $\{d_w\}$ is bounded, we have that

$$\|A^Td_0 + c\|^2 \leq \xi^2. \tag{3.13}$$

Hence $d_0 \in \mathcal{F}_0$.

Since for all global solution d_0^* of P_0, $d_0^* \notin \mathcal{F}_1$, and $d_0 \in \mathcal{F}_1$,

$$\Phi(d_0^*) < \Phi(d_0). \tag{3.14}$$

However, in the neighborhood of d_0^*, for each $w > 0$ sufficiently small, there exists a point $\tilde{d}_w \in \mathcal{F}_w$, and $d_w^* \to d_0^*$. Now, $\tilde{d}_w, d_w \in \mathcal{F}_w$ with $d_w \in S(w)$ being global minimizer of problem P_w,

$$\Phi(d_w^*) \geq \Phi(d_w), \tag{3.15}$$

taking $w \to 0+$, we deduce an inequality (3.14) with an opposite inequality sign.

4. Relations Between the CDT Problem and P_w

This section we show the relations between the CDT problem and problem P_w. Recalling that we divide into three cases for trust region subproblem, one may think problem P_w will in hard case 2) of the three cases if the CDT problem has no global solution with the Hessian of Lagrangian positive semi-definite. We state the result as theorem 4.1.

Theorem 4.1. *If for any $w \in [0,1]$, no global solution of problem P_w is feasible to the CDT problem, there exists a $w \in (0,1)$ such that the global solutions of problem P_w is in the case 2) of hard case.*

Note that the result means the global solution set of problem P_w includes two points, furthermore, each of these two points is feasible to either one of $\mathcal{F}_0, \mathcal{F}_1$ respectively.

Proof. If, for any $w \in [0,1]$ there exists no global solution of P_w feasible to the CDT problem, B can not be positive semi-definite. It is also easy to see that,

$$S(0) \cap \mathcal{F}_1 = \emptyset \text{ and } S(1) \cap \mathcal{F}_0 = \emptyset, \tag{4.1}$$

otherwise we obtain the global solution $d_0 \in S(0)$ (or $d_1 \in S(1)$) which is feasible to the CDT problem and hence it is the desired global solution of the CDT subproblem with the Hessian of Lagrangian positive semi-definite(another multiplier is 0).

Since for all $d_1 \in S(1)$, $\|A^T d_1 + c\| > \xi$, there exists a $\bar{w}$, $0 < \bar{w} < 1$ with $1 - \bar{w}$ sufficiently small, such that for all $d_{\bar{w}} \in S(\bar{w})$, we also have $\|A^T d_{\bar{w}} + c\| > \xi$, hence $\|d_{\bar{w}}\| < \Delta$. We also have, for all $d_0 \in S(0)$, $\|d_0\| > \Delta$.

We claim the following result:

$$S(w) \cap \{d \mid \|d\| = \Delta\} = \emptyset, \ \forall w \in (0, \bar{w}). \tag{4.2}$$

If (4.2) is not true, then there exists $w \in (0, \bar{w})$ and $d_w \in S(w)$ such that $\|d_w\| = \Delta$, which implies $d_w \in \mathcal{F}$. Thus d_w is also a global minimizer of the original CDT problem, and the corresponding Hessian of Lagrangian $H(\lambda, \mu) = B + \tau^*(wI + (1-w)AA^T)$ is positive semi-definite, which contradicts our assumptions.

From Lemma 3.1, (4.2) and the facts that

$$\|d\| > \Delta, \quad \forall d \in S(\bar{w}), \tag{4.3}$$

$$\|d\| < \Delta, \quad \forall d \in S(0), \tag{4.4}$$

there is at least one $\hat{w} \in (0, \bar{w})$ such that there exist $d_1, d_2 \in S(\hat{w})$ satisfying

$$\|d_1\| > \Delta, \quad \|d_2\| < \Delta. \tag{4.5}$$

(4.5), (4.2) and Theorem 2.2 imply that the global solutions of problem $P_{\hat{w}}$ is in the case 2) of the hard case.

This lemma does not exclude the following circumstance: there exists $w \in (0,1)$ such that case 2) of hard case holds and for all $d \in S(w)$, $\|d\| < \Delta$ or $\|d\| > \Delta$.

Visually say, $d_w \in S(w) \subset \mathcal{F}_1 \cup \mathcal{F}_0$ is a "curve" when w variants form 0 to 1 continuously, if there is a $d_w \in \mathcal{F}$, then we get the global solution of the CDT problem. Since $d_0 \in \mathcal{F}_0 \setminus \mathcal{F}_1$, $d_1 \in \mathcal{F}_1 \setminus \mathcal{F}_0$, and for all $w \in (0,1)$, $d_w \in \mathcal{F}_1 \cup \mathcal{F}_0$, there will be a "jump" form $\mathcal{F}_0 \setminus \mathcal{F}_1$ to $\mathcal{F}_1 \setminus \mathcal{F}_0$, if there is no global solution of the CDT problem with the Hessian of Lagrangian positive semi-definite. This is what theorem 4.1 illustrates.

Corollary 4.1. *If there is no global solution of the CDT problem with the Hessian of Lagrangian positive semi-definite, then there exists at most one $w \in (0,1)$ such that case 2) of hard case holds for problem P_w and $d_1, d_2 \in S(w)$ satisfying (4.5).*

Proof. By contradiction. Suppose that there exist $w_1, w_2 \in (0,1)$, $w_1 < w_2$ such that $S(w_1) = \{p_1, p_2\}$, $S(w_2) = \{p_3, p_4\}$ satisfying the following equation:

$$w_i(\|p_j\|^2 - \Delta^2) + (1 - w_i)(\|A^T p_j + c\|^2 - \xi^2) = 0, \tag{4.6}$$

where $j \in \{1,2\}$ if $i = 1$, $j \in \{3,4\}$ if $i = 2$, and

$$\|p_i\| < \Delta, \ \|A^T p_i + c\| > \xi, \quad i = 2, 4, \tag{4.7}$$

$$\|p_i\| > \Delta, \ \|A^T p_i + c\| < \xi, \quad i = 1, 3. \tag{4.8}$$

Since that

$$\begin{aligned} 0 &= w_2(\|p_3\|^2 - \Delta^2) + (1 - w_2)(\|A^T p_3 + c\|^2 - \xi^2) \\ &= \frac{1 - w_2}{1 - w_1}\left(w_1(\|p_3\|^2 - \Delta^2) + (1 - w_1)(\|A^T p_3 + c\|^2 - \xi^2)\right) \\ &\quad + \frac{w_2 - w_1}{1 - w_1}(\|p_3\|^2 - \Delta^2), \end{aligned} \tag{4.9}$$

it follows from $0 < w_1 < w_2 < 1$ and $\|p_3\| > \Delta$ that

$$w_1(\|p_3\|^2 - \Delta^2) + (1 - w_1)(\|A^T p_3 + c\|^2 - \xi^2) < 0, \tag{4.10}$$

which shows $p_3 \in \mathcal{F}_{w_1}$. Remember that $p_2 \in S(w_1) \subset \mathcal{F}_{w_1}$, thus $\Phi(p_2) \leq \Phi(p_3)$. Similarly, we can prove that $p_2 \in \mathcal{F}_{w_2}$ and $\Phi(p_3) \leq \Phi(p_2)$. Then $\Phi(p_3) = \Phi(p_2)$, p_3 is a global minimizer of problem P_{w_1}, which contradicts to $S(w_1) = \{p_1, p_2\}$. The contradiction completes our proof.

Since the "jump" is unique, we would ask what kind of properties the jump has. In the following two sections, we give the geometry meaning of the jump, and study the maxima of dual function of the CDT problem by the jump. However, if let d_{appro} be a feasible point to the CDT problem which lies on the line $d = d_1 + t(d_2 - d_1)$, where d_1, d_2 are defined in (4.5), one natural question is whether d_{appro} is a good approximation to the global solution of original the CDT problem? Unfortunately we are not able to answer this question yet.

Theorem 4.2. *If there exists a global solution d^* of the CDT problem with the corresponding Hessian of Lagrangian positive semi-definite, then there exists a $w \in [0,1]$ such that d^* is also a global solution of problem P_w.*

To prove this theorem, first we introduce optimality conditions for problem (3.12).

Theorem 4.3. *Assume that problem (3.12) has finite global optimal value. A feasible point d^* is a global solution of problem (3.12) if and only if there exists $\mu^* \geq 0$ such that*

$$(B + \mu^* AA^T)d^* = -(g + \mu^* Ac), \tag{4.11}$$

and

$$\mu^*(\|A^T d^* + c\| - \xi) = 0 \tag{4.12}$$

with the Hessian of Lagrangian $B + \mu^* AA^T$ positive semidefinite.

Proof. If there exists μ^* such that (4.11) and (4.12) hold with $B + \mu^* AA^T$ positive semidefinite, then d^* is a global minimizer of

$$\widehat{\Phi}(d) = \frac{1}{2} d^T(B + \mu^* AA^T)d + (g + \mu^* Ac). \tag{4.13}$$

Hence

$$\widehat{\Phi}(d^*) \leq \widehat{\Phi}(d), \tag{4.14}$$

holds for all $d \in \mathcal{R}^n$ and

$$\begin{aligned} \Phi(d^*) \quad &= \widehat{\Phi}(d^*) - \tfrac{\mu^*}{2}(\|A^T d^* + c\|^2 - \|c\|^2) \\ &\leq \widehat{\Phi}(d) - \tfrac{\mu^*}{2}(\|A^T d + c\|^2 - \|c\|^2) \\ &= \Phi(d), \end{aligned} \tag{4.15}$$

holds for all d satisfying $\|A^T d + c\| \leq \xi$. Therefore, d^* is a global solution of problem (3.12).

On the other hand, if d^* is a global minimizer of problem (3.12), by K-T theory, (4.11) and (4.12) hold. If $\|A^T d^* + c\| < \xi$, $\mu^* = 0$ and $Bd^* = -g$, hence B is positive semidefinite. If $\|A^T d^* + c\| = \xi$, the second order necessary condition implies that for all v satisfying $v^T y^* = 0$, we have,

$$v^T(B + \mu^* AA^T)v \geq 0, \tag{4.16}$$

where $y^* = A(A^T d^* + c)$. If $v^T y^* = v^T A(A^T d^* + c) \neq 0$, it is easy to see that $A^T v \neq 0$. Since for all d satisfying $\|A^T d + c\| = \xi$, (4.15) holds, substitute $g + \mu^* Ac$ by (4.11), we have

$$\frac{1}{2}(d^* - d)^T(B + \mu^* AA^T)(d^* - d) \geq 0 \tag{4.17}$$

holds for all d satisfying $\|A^T d + c\| = \xi$. Let $d = d^* - 2tv$, where

$$t = \frac{-2v^T y^*}{\|Av\|^2}, \tag{4.18}$$

which implies that $\|A^T d + c\| = \xi$ and (4.17) means by simple calculations that

$$v^T(B + \mu^* AA^T)v \geq 0. \tag{4.19}$$

(4.16) and (4.19) complete our proof.

Now, we are ready to prove theorem 4.2.

Proof of theorem 4.2. Suppose the corresponding multiplier of d^* is (λ^*, μ^*). By the assumption, $H(\lambda^*, \mu^*)$ is positive semi-definite.

If $\lambda^* = 0$ and $\mu^* \geq 0$, then (4.11) and (4.12) hold with $B + \mu^* AA^T$ positive semidefinite, then d^* is a global solution to both the CDT problem and problem P_0 by theorem 4.4. If $\lambda^* \geq 0$ and $\mu^* = 0$, Analogous considerations on (2.1)–(2.2) and theorem 2.1, we have the same result.

If $\lambda^* > 0$ and $\mu^* > 0$, then both constraints of the CDT problem are active at d^*, then the constraints of problem P_w is active at d^* for all $w \in [0, 1]$. Consider problem $P_{\frac{\lambda^*}{\lambda^*+\mu^*}}$, the point d^* and the Lagrangian multiplier $\tau^* = \lambda^* + \mu^*$. It can be seen that d^* is a global solution of problem $P_{\frac{\lambda^*}{\lambda^*+\mu^*}}$ with Lagrangian multiplier $\tau^* = \lambda^* + \mu^*$ and the Hessian of Lagrangian $B + \lambda^* I + \mu^* AA^T$ positive semidefinite by the equivalent transformed problem (3.4)–(3.5).

5. Geometry Meaning of Jump

If there exists a global solution of the CDT problem with the Hessian of Lagrangian positive semi-definite, a global minimizer can be found by the following two approaches:

1) to search the global solution of the CDT problem in the region Ω_0 by the dual algorithm, for example, see Yuan (1991);

2) to find a global solution of the scaled problem P_w such that the solution is feasible to the CDT problem.

However, if the above assumption fails, process 1) terminates at the maxima of dual function $\Psi(\lambda, \mu)$, while process 2) will find a "jump", i.e., case 2) of hard case of scaled problem occurs with two solutions satisfying (4.5) respectively. In this section we consider the relation between them.

First, we discuss the relations between the multiplier of problem P_w and the multipliers of the CDT problem. It can be shown that if there exists (τ_w, d_w) satisfying

$$\left(B + \tau_w(wI + (1-w)AA^T)\right) d_w = -(g + \tau_w(1-w)Ac) \tag{5.1}$$

with $d_w \in S(w)$, i.e., the solution of problem P_w, then the triple $(\lambda, \mu, d(\lambda, \mu)) = (\tau_w w, \tau_w(1 - w), d_w)$ satisfies the first equation of KKT system of the CDT problem with the Hessian of Lagrangian positive semi-definite. On the other hand, if there exists $(\lambda, \mu, d) \neq 0$ satisfying

$$(B + \lambda I + \mu AA^T)d = -(g + \mu Ac) \tag{5.2}$$

with $H(\lambda, \mu)$ positive semi-definite and defect 1, then $(\lambda + \mu, d)$ satisfies the first equation of KKT system of problem P_w with $w = \frac{\lambda}{\lambda+\mu}$.

In the following, we consider the multipliers which are the maxima of the dual function (1.10) on the region Ω_0. First we give a definition as follow.

Definition 5.1. *We call that the nonzero multipliers (λ, μ) have property $\mathcal{J}$ of problem (1.1)–(1.3) if*
i) *$H(\lambda, \mu)$ is positive semidefinite with defect 1;*
ii) *the scaled problem $P_{\frac{\lambda}{\lambda+\mu}}$ is in hard case 2) and its global solutions satisfy inequalities (4.5).*

Thus by definition 5.1, multipliers (λ, μ) having property $\mathcal{J}$ iff the scaled problem $P_{\frac{\lambda}{\lambda+\mu}}$ of problem P is in the case 2 of hard case. Thus, for all (λ, μ) satisfies property $\mathcal{J}$, $w = \frac{\lambda}{\lambda+\mu}$ has the same value.

We give the main result of this section as follow:

Theorem 5.1. *If there is no global solution of the CDT problem with the Hessian of Lagrangian positive semi-definite, the multipliers (λ, μ) satisfying property $\mathcal{J}$ is identical to the multipliers being a maxima of dual function in region Ω_0.*

Before we prove Theorem 5.2, we give some lemmas. Lemma 5.3 is about the invariant of property $\mathcal{J}$.

Lemma 5.1. *The multipliers (λ, μ) satisfies property $\mathcal{J}$ of problem (1.1)–(1.3) if and only if, for all $(\delta\lambda, \delta\mu) \in [0, \lambda) \times [0, \mu)$, the multipliers $(\lambda - \delta\lambda, \mu - \delta\mu)$ satisfies property $\mathcal{J}$ of problem $\widehat{P}$:*

$$\min_{\hat{d} \in \mathcal{R}^n} \widehat{\Phi}(\hat{d}) = \frac{1}{2}\hat{d}^T \widehat{B} \hat{d} + \hat{g}^T \hat{d} \tag{5.3}$$

subject to (1.2)–(1.3) for $\hat{d}$, where $\widehat{B} = B + \delta\lambda I + \delta\mu A A^T$ and $\hat{g} = g + \delta\mu Ac$.

Proof. Except for a constant, the objective function $\hat{\Phi}(d)$ is the sum of the original objective function $\Phi(d)$ and a penalty term $\frac{1}{2}(\delta\lambda \|d\|^2 + \delta\mu ||A^T d + c||^2)$. Moreover, there exists certain relations between counterparts of both two problems, P and $\widehat{P}$, for example, the Hessian of Lagrangian, the regions where the Hessian of Lagrangian has one negative eigenvalue, etc, for more detail, see Chen and Yuan (1998).

The Hessian $H(\lambda, \mu)$ retains part i) of property $\mathcal{J}$ since

$$H(\lambda + \delta\lambda, \mu + \delta\mu) = \widehat{H}(\lambda, \mu), \tag{5.4}$$

where $\widehat{H}$ is the Hessian of Lagrangian of problem $\widehat{P}$. It means they possess the same null space. Since we also have

$$g + (\mu + \delta\mu) Ac = \hat{g} + \mu Ac, \tag{5.5}$$

and inequalities (4.5) are only related to original problem, except for the null space step which is invariant by (5.4), part ii) of property $\mathcal{J}$ is satisfied for both problem P and $\widehat{P}$ due to the fact that

$$d(\lambda + \delta\lambda, \mu + \delta\mu) = \hat{d}(\lambda, \mu), \tag{5.6}$$

where $d(\lambda, \mu)$ is given by

$$d(\lambda, \mu) = \begin{cases} \text{undefined}, & \text{if (5.2) is inconsistent}, \\ -H(\lambda, \mu)^+ (g + \mu Ac), & \text{otherwise} \end{cases} \tag{5.7}$$

and $\hat{d}(\lambda, \mu)$ is defined similarly. Hence (λ, μ) satisfies property $\mathcal{J}$ of P iff it satisfies property $\mathcal{J}$ of $\widehat{P}$.

We restate the definition of singular line on the dual plane $\mathcal{R}^2_+$, also see Chen and Yuan (1998). The line $\lambda = \lambda_s \geq 0$ is called the singular line if

$$det(H(\lambda_s, \mu)) \equiv 0, \text{ for all } \mu \geq \mu_s, \tag{5.8}$$

with $H(\lambda_s, \mu)$ positive semidefinite, where μ_s is a scale. And we also define μ_c as

$$\mu_c = arg\ \min\{\mu \geq 0 \mid H(\lambda_s, \mu) \text{ is positive semidefinite}\}. \tag{5.9}$$

We prove Theorem 5.2 by considering three cases:

- case 1: the multipliers $(\lambda_{max}, \mu_{max})$ being a maxima of dual function is not on the singular line, i.e., $\lambda \neq \lambda_s$;
- case 2: $\lambda_{max} = \lambda_s$ and $\mu_{max} > \mu_c$;
- case 3: $\lambda_{max} = \lambda_s$ and $\mu_{max} = \mu_c$.

The following lemma deals with case 1.

Lemma 5.2. *If there is no global solutions of problem (1.1)–(1.3) with the Hessian of Lagrangian positive semi-definite, then the multipliers (λ, μ) not in the singular line satisfies property $\mathcal{J}$ if and only if*

$$(\lambda, \mu) = arg \max_{(\lambda, \mu) \in \Omega_0} \Psi(\lambda, \mu). \tag{5.10}$$

Proof. Suppose that (λ_1, μ_1) satisfying property $\mathcal{J}$ is not in the singular line. Let
$$(\lambda_1, \mu_1) \neq (\lambda_2, \mu_2) = arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda, \mu). \tag{5.11}$$
It is easy to see that $(\lambda_1, \mu_1), (\lambda_2, \mu_2) \in \partial\Omega_0$ because there is no global solutions of problem (1.1)–(1.3) with the Hessian positive semi-definite. By the structure of dual plane, please see Chen and Yuan (1998) for more detail, for any two points $(\lambda_i, \mu_i) \in \partial\Omega_0, i = 1, 2$, we have
$$(\lambda_1 - \lambda_2)(\mu_1 - \mu_2) \leq 0. \tag{5.12}$$
If $\lambda_1 = \lambda_2$, our assumptions imply that $det(H(\lambda_1, \mu)) = 0$ for all μ between μ_1 and μ_2. Hence, $\lambda_1 = \lambda_2$ is in the singular line, which is a contradiction. If $\mu_1 = \mu_2$, then either $\lambda_1 > \lambda_2$ or $\lambda_2 > \lambda_1$. If $\lambda_1 > \lambda_2$, one can see that $H(\lambda_1, \mu_1)$ is positive definite, which contradicts the fact that (λ_1, μ_1) has property $\mathcal{J}$. On the other hand, if $\lambda_2 > \lambda_1$, $H(\lambda_2, \mu_2)$ is positive definite, which contradicts the fact that there is no global solutions of problem (1.1)–(1.3) with a positive semidefinite Hessian of Lagrangian. Thus we have proved that
$$(\lambda_1 - \lambda_2)(\mu_1 - \mu_2) \neq 0, \tag{5.13}$$
which, because of (5.12), implies that
$$\lambda_1 < \lambda_2 \qquad \text{or} \qquad \mu_1 < \mu_2. \tag{5.14}$$
Without loss of generality, we suppose $\lambda_1 < \lambda_2$. Consider the problem $\widehat{P}$ with $\delta\lambda = \frac{1}{2}(\lambda_1 + \lambda_2)$ and $\delta\mu = 0$. Then problem $\widehat{P}$ has no point (λ_j, μ_j) having property $\mathcal{J}$, otherwise, problem P has two points, (λ_1, μ_1) and $(\lambda_j + \delta\lambda, \mu_j + \delta\mu)$ due to lemma 5.3, having property $\mathcal{J}$ which contradicts Corollary 4.2. By Theorem 4.1, problem $\widehat{P}$ has a global solution with the Hessian of Lagrangian positive semi-definite, which shows its corresponding multiplier is $(\lambda_2 - \delta\lambda, \mu_2)$ implied by the assumption (5.11). Then, problem P also has a global solution with corresponding multipliers (λ_2, μ_2) implied by the theorem 4.4 in Chen and Yuan (1998), since $\delta\mu = 0$ and $\lambda_2 - \delta\lambda \neq 0$. This contradicts to the assumption of this lemma.

Now we consider case 2.

Lemma 5.3. *Let (λ_s, μ_{max}) is a maxima of the dual function on Ω_0. If $\mu_{max} > \mu_c$, then there is a global solution of the CDT subproblem with the Hessian positive semidefinite, and (λ_s, μ_{max}) is the corresponding multipliers.*

Proof. Since (λ_s, μ_{max}) is a maxima of the dual function onΩ_0 with $\mu_{max} > \mu_c$,
$$\|d_{max}\| \leq \Delta \qquad \text{and} \qquad \|A^T d_{max} + c\| = \xi, \tag{5.15}$$
where, $d_{max} = d(\lambda_s, \mu_{max})$.

Due to (5.8), there exists a vector $d \in R^n$ such that
$$A^T d = 0 \tag{5.16}$$
and
$$(B + \lambda_s I + \mu_{max} AA^T)d = 0. \tag{5.17}$$
Because (5.15) and (5.16), there exists a $t \in R$ such that
$$\|d_{max} + td\| = \Delta \quad \text{and} \quad \|A^T(d_{max} + td) + c\| = \xi. \tag{5.18}$$
Now it is easy to see that $d_{max} + td$ is a global solution of the CDT subproblem with the Hessian of Lagrangian $H(\lambda_s, \mu_{max})$ positive semidefinite.

Lemma 5.5 shows that if the maxima of the dual function on Ω_0 locates on the singular line with the Hessian defect 1 and $\mu_{max} > \mu_c$, there exists a global solution of the CDT problem with a positive semidefinite Hessian. The following result shows that any point on the singular line does not satisfy property $\mathcal{J}$.

Lemma 5.4. *The multipliers $(\lambda_s, \mu) \in \partial\Omega_0$ does not satisfy property $\mathcal{J}$ for any $\mu > \mu_c$.*

Proof. If (λ_s, μ) on the singular line satisfies property $\mathcal{J}$, the Hessian of Lagrangian with multipliers on singular line must have one multiple zero eigenvalue.

If (λ_s, μ) satisfies property $\mathcal{J}$, $H(\lambda_s, \mu)$ has one multiple zero eigenvalue. For any $v \in \mathcal{N}(H(\lambda_s, \mu))$, we have $A^T v = 0$. Hence, for any fixed $\lambda = \lambda_s$, $\mu > \mu_c$, for all d satisfying (5.2),
$$\|A^T d + c\| - \xi = const, \tag{5.19}$$
thus (4.5) can not be satisfied at the same time.

Finally we study the case 3 when $\lambda_{max} = \lambda_s$ and $\mu_{max} = \mu_c$.

Lemma 5.5. *If $\lambda_{max} = \lambda_s$ and $\mu_{max} = \mu_c$, $(\lambda_{max}, \mu_{max})$ is multipliers of a global solution of the CDT problem.*

Proof. If $\mu_c > 0$, $H(\lambda_s, \mu_c)$ has at least two multiples of zero eigenvalue and (λ_s, μ_c) does not satisfies property $\mathcal{J}$. Hence it follows by Lemma 5.6 there is no multipliers on the singular line satisfying property $\mathcal{J}$. By lemma 5.4, if the assumption of Theorem 5.2 holds, it is also impossible for any multipliers not on the singular line to satisfy property $\mathcal{J}$. Then there is no multipliers on the dual plane satisfying property $\mathcal{J}$, and thus the CDT problem has a global solution with the Hessian of Lagrangian positive semi-definite.

If $\mu_c = 0$, for any (λ, μ) with $\lambda > \lambda_s$ and $\mu \geq 0$, $H(\lambda, \mu)$ is positive definite. Hence property $\mathcal{J}$ can only be satisfied in $(\lambda, \mu) = (\lambda_s, 0)$. In this case, problem P_w is actually P_1 with the global solutions satisfying (4.5). Let

$$d(\theta) = d_1 + \theta(d_2 - d_1), \tag{5.20}$$

θ can be chosen such that $\|d(\theta)\| = \Delta$. Since $\|A^T d_i + c\| \leq \xi$, $i = 1, 2$, $\|A^T d(\theta) + c\| \leq \xi$. Therefore, $d(\theta)$ is a global solution of the CDT problem with the Hessian of Lagrangian positive semi-definite.

From Lemmas 5.4, 5.5, 5.6 and 5.7, we see that Theorem 5.2 is true. We present a new proof of Theorem 4.1 in Chen and Yuan (1998).

Corollary 5.1. *If the right hand side of (5.11) is a segment, then there exists a global solution of the CDT subproblem with the Hessian positive semi-definite.*

Proof. Suppose that the CDT subproblem has no global solution with the Hessian positive semi-definite. It follows from Theorem 5.2 that each $(\lambda, \mu) \in T$ satisfies property $\mathcal{J}$, where T is defined by

$$(\lambda, \mu) \in T = arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda, \mu). \tag{5.21}$$

If T is a segment, then $w = \frac{\lambda}{\lambda+\mu}, ((\lambda, \mu) \in T)$ is also a segment, which contradicts Corollary 4.2.

6. Further Study of Dual Maxima

We will extend the result about the location of global solutions of the CDT subproblem by the existence and uniqueness of multipliers satisfying property $\mathcal{J}$. For more detail about the location of global solutions, please see Chen and Yuan (1998).

Theorem 6.1. *If $(\lambda, \mu) \in \partial\Omega_0$ satisfies (5.21), and if either one of (a) $\lambda \cdot \mu = 0$, and (b) $H(\lambda, \mu)$ has at least two multipliers of zero eigenvalue holds, then (λ, μ) corresponds a global solution of the CDT subproblem with the Hessian $H(\lambda, \mu)$ positive semidefinite.*

Proof. We need to prove that there exists no multipliers satisfying property $\mathcal{J}$, which implies, by theorem 4.1, there exists a global solution with the Hessian positive semidefinite and its corresponding multipliers must be (λ, μ).

Part (a) can be shown similarly as Lemma 5.7.

Suppose that (b) holds. Assume that there is no global solution with the Hessian positive semidefinite, (λ, μ) satisfies property $\mathcal{J}$ since (λ, μ) is not on singular line. However, the first condition of property $\mathcal{J}$ fails for (λ, μ), which is a contradiction.

Now, we can conclude the result on the location of global solutions as following:

Theorem 6.2. *If for any $(\lambda, \mu) \in T = \{(\lambda, \mu)\}$ satisfies (5.21) and either one of the following conditions:*
(a) *there exists $(\lambda, \mu) \in T \cap int\Omega_0$;*
(b) *$T = \{(\lambda, \mu)\}$ is not a singleton;*
(c) *there exists $(\lambda, \mu) \in T$ such that $\lambda \cdot \mu = 0$;*
(d) *there exists $(\lambda, \mu) \in T$ such that the Hessian $H(\lambda, \mu)$ has two multipliers zero eigenvalue;*
(e) *there exists $(\lambda, \mu) \in T$ on the singular line,*
then there is a global solution of the CDT subproblem with the Hessian $H(\lambda, \mu)$ positive semi-definite.

Recall the definitions of Ω_1, Ω_{1k}, $l_{\lambda k}$ and etc, on dual plane, please see Chen and Yuan (1998), from Theorem 6.2, we can get the following result.

Corollary 6.1. *For $(\lambda, \mu) \in T = \{(\lambda, \mu)\}$ satisfies (5.21), (λ, μ) does not correspond a global solution only if*
(a) *T is a singleton and (λ, μ) is not on the singular line;*
(b) *there exists a certain connected branch Ω_{1k}, such that*

$$l_{\lambda k} < \lambda < u_{\lambda k} \quad \text{and} \quad l_{\mu k} < \mu < u_{\mu k}. \tag{6.1}$$

References

[1] M.R. Celis, J.E. Dennis, R.A. Tapia, A trust region algorithm for nonlinear equality constrained optimization, in: P.T. Boggs, R.H. Byrd and R.B. Schnabel,eds., Numerical Optimization, (SIAM, Philadelphia, 1985), pp.71-82

[2] X.D. Chen, Y.X. Yuan, On local solutions of the CDT subproblem, Report ICM-98-8, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, China, submitted to SIAM J.Optimization, 1998.

[3] M.M. El-Alem, R.A. Tapia, Numerical experience with a polyhedral-norm CDT trust-region algorithm, *Journal of Optimization Theory and Applications,* **85** (1995), 575-591.

[4] M. Fu, Z. Luo, Y. Ye, Approximation algorithms for quadratic programming, July, 1996.

[5] D.M. Gay, Computing optimal local constrained step, *SIAM J. Sci. Stat. Comp.,* **2** (1981), 186-197.

[6] J.J. Moré, D.C. Sorensen, Computing a trust region step, *SIAM J. Sci. Stat. Comp,* **4** (1983), 553-572.

[7] M.J.D. Powell, Y. Yuan, A trust region algorithm for equality constrained optimization, *Math. Prog.,* **49** (1991), 189-211.

[8] D.C. Sorensen, Newton's method with a model trust region modification, *SIAM J.Numer.Anal.,* **19** (1982), 409-426.

[9] Y. Yuan, On a subproblem of trust region algorithms for constrained optimization, *Math. Prog.,* **47** (1990), 53-63.

[10] Y. Yuan, A dual algorithm for minimizing a quadratic function with two quadratic constraints, *J. Comp. Math.,* **9** (1991), 348-359.

[11] Y. Zhang, Computing a Celis-Dennis-Tapia trust-region step for equality constrained optimization, *Math. Prog.,* **55** (1992), 109-124.