Journal of Computational Mathematics, Vol.23, No.5, 2005, 463–478.

COMPUTE A CELIS-DENNIS-TAPIA STEP *1)

Gai-di Li
(*College of Applied Sciences, Beijing University of Technology, Beijing 100022, China*)
(*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Ya-xiang Yuan
(*LSEC, ICMSEC, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*)

Abstract

In this paper, we present an algorithm for the CDT subproblem. This problem stems from computing a trust region step of an algorithm, which was first proposed by Celis, Dennis and Tapia for equality constrained optimization. Our algorithm considers general case of the CDT subproblem, convergence of the algorithm is proved. Numerical examples are also provided.

Mathematics subject classification: 65K10, 90C30.
Key words: The CDT subproblem, Local solution, Global solution, Dual function.

1. Introduction

The CDT subproblem has the following form,

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

$$\text{s.t.} \quad \|d\| \le \Delta,\ \|A^Td + c\| \le \xi, \tag{1.2}$$

where $g \in R^n$, $A \in R^{n\times m}$, $c \in R^m$, $\Delta > 0$, $\xi \ge 0$, and $B \in R^{n\times n}$ is a symmetric matrix. Throughout this paper, the norm $||\cdot||$ is the Euclidean norm. For convenience, we denote by $\mathcal{F}$ the feasible region of the CDT subproblem, namely,

$$\mathcal{F} = \{d \,|\, ||d|| \le \Delta, ||A^Td + c|| \le \xi\}. \tag{1.3}$$

As an important application, the CDT subproblem is a subproblem of some trust region algorithms for nonlinear programming, which was given by Celis, Dennis & Tapia[1] and Powell & Yuan[10], whose superlinear convergence property is obtained under certain conditions.

The properties of the CDT subproblem have been studied by many researchers, see Yuan[12], Peng & Yuan[9] and Chen & Yuan[2, 3, 4] etc. With some additional assumptions, some algorithms have been presented. For example, under the assumption that B is positive definite, two different algorithms have been proposed by Yuan[13] and Zhang[14] respectively. In this paper, we present an algorithm for solving problem (1.1)–(1.2) for general symmetric matrix B. We also assume that $\mathcal{F}$ has strict interior points.

The paper is organized as follows. In the next section, we state some known results which we will use in this paper. In section 3, we consider dual function and give some useful results.

* Received July 5, 2004; final revised October 12, 2004.
1) Supported by Chinese NSF grant 10231060 and Beijing University of Technology research grants for young teachers.

In section 4, the algorithm is presented. In section 5, convergence properties are analyzed. And in the last two section, we give numerical experiments.

2. Some Basic Results

In this section, we restate some fundamental results of the CDT subproblem.

Firstly, we define some notations as follows,

$$\mathcal{R}_+^2 = \{(\lambda,\mu) \mid \lambda \ge 0,\ \mu \ge 0\}, \tag{2.1}$$

$$\Omega_0 = \{(\lambda,\mu) \in \mathcal{R}_+^2 \mid H(\lambda,\mu) \text{ is positive semi-definite}\}, \tag{2.2}$$

$$\Omega_1 = \{(\lambda,\mu) \in \mathcal{R}_+^2 \mid H(\lambda,\mu) \text{ has one negative eigenvalue}\}, \tag{2.3}$$

where $H(\lambda,\mu) = B + \lambda I + \mu AA^T$. For the dual variables $\lambda \ge 0, \mu \ge 0$, when $g + \mu Ac \in \mathcal{R}(H(\lambda,\mu))$, we also define the vector

$$d(\lambda,\mu) = -H(\lambda,\mu)^+(g + \mu Ac), \tag{2.4}$$

and the Lagrangian dual function

$$\Psi(\lambda,\mu) = \Phi(d(\lambda,\mu)) + \frac{\lambda}{2}(||d(\lambda,\mu)||^2 - \Delta^2) + \frac{\mu}{2}(||A^T d(\lambda,\mu) + c||^2 - \xi^2). \tag{2.5}$$

The optimal conditions for the CDT subproblem were first proved by Yuan[12].

Theorem 2.1. *Let d^* be a global solution of the problem (1.1)–(1.2). Then there exist $\lambda^*, \mu^* \ge 0$ such that*

$$(B + \lambda^* I + \mu^* AA^T)d^* = -(g + \mu^* Ac), \tag{2.6}$$

$$\lambda^*(\Delta - ||d^*||) = 0, \quad \mu^*(\xi - ||A^T d^* + c||) = 0. \tag{2.7}$$

Furthermore, the matrix

$$H(\lambda^*,\mu^*) = B + \lambda^* I + \mu^* AA^T \tag{2.8}$$

has at most one negative eigenvalue if the multiplier λ^*, μ^* are unique.

In addition, Yuan[12] also showed that we could always find the global solution d^* with $(\lambda^*,\mu^*) \in \Omega_0 \cup \Omega_1$. From these results, we can try to get the global solution of the CDT subproblem by searching the dual variables (λ^*,μ^*) in $\Omega_0 \cup \Omega_1$. Further, Chen & Yuan[2] gave the criterion that the CDT subproblem has no global solution d^* with $(\lambda^*,\mu^*) \in \Omega_0$, which can be presented as follows.

Theorem 2.2. *If there is no global solution d^* of the CDT subproblem with the corresponding $H(\lambda^*,\mu^*)$ positive semi-definite, then the maxima (λ_+,μ_+) of dual function $\Psi(\lambda,\mu)$ in the region Ω_0 satisfy*

i) *$H(\lambda_+,\mu_+)$ is positive semi-definite with defect 1;*

ii)

$$(||d(\lambda_+,\mu_+) + \tilde\tau_+\tilde u|| - \Delta)(||d(\lambda_+,\mu_+) + \tilde\tau_-\tilde u|| - \Delta) < 0, \tag{2.9}$$

where $d(\lambda_+,\mu_+)$ is defined by (2.4), $\tilde u$ satisfies $||\tilde u|| = 1$ and $H(\lambda_+,\mu_+)\tilde u = 0$, and $\tilde\tau_\pm$ are respectively given by

$$\lambda_+(||d(\lambda_+,\mu_+) + \tilde\tau\tilde u||^2 - \Delta^2) + \mu_+(||A^T(d(\lambda_+,\mu_+) + \tilde\tau\tilde u) + c||^2 - \xi^2) = 0. \tag{2.10}$$

It is well known that a KKT point d^* is the global solution of the CDT subproblem if H is positive semi-definite. About the KKT points d^* with $(\lambda^*, \mu^*) \in \Omega_1$, Li [7] gave the local optimal conditions when $\lambda^*\mu^* > 0$; If $\lambda^*\mu^* = 0$, the optimal conditions are reduced to the conditions when this point is a local solution of a simpler problem, which is easy to verify. Furthermore, Li [7] showed that all KKT points in Ω_1 were local optimal solutions.

Based on these theories, we present an dual algorithm by restricting $(\lambda, \mu) \in \Omega_0$ at the first stage; if we don't find the global solution in Ω_0, we can enter the second stage by the criterion of theorem 2.2, namely, we extend the search region from Ω_0 to $\Omega_0 \bigcup \Omega_1$.

In the end of this section, we give the assumption as follows.

Assumption 2.3.

a) *Assume that there is at least one active constraint at the solution d^* of the CDT subproblem (1.1)-(1.2).*

b) *Assume that*

$$\min_{(\lambda,\mu)\in L_0(\lambda_0,\mu_0)\cup\Omega_1} \{||d(\lambda,\mu)||, ||A^T d(\lambda,\mu) + c||\} \geq c_0 \tag{2.11}$$

holds, where $d(\lambda, \mu)$ is the same as (2.4), and $L_0(\lambda_0, \mu_0)((\lambda_0, \mu_0) \in \Omega_0)$ is defined by

$$L_0(\lambda_0, \mu_0) = \{(\lambda, \mu) \,|\, \Psi(\lambda, \mu) > \Psi(\lambda_0, \mu_0), (\lambda, \mu) \in \Omega_0\}.$$

3. Dual Function

It is easy to see that dual function for the CDT subproblem (1.1)-(1.2) can be presented as

$$\Psi(\lambda, \mu) = \begin{cases} \text{undefined}, \quad \text{if } H(\lambda,\mu)d(\lambda,\mu) = -(g + \mu Ac) \text{ is inconsistent}, \\ \Phi(d(\lambda,\mu)) + \frac{\lambda}{2}(||d(\lambda,\mu)||^2 - \Delta^2) + \frac{\mu}{2}(||A^T d(\lambda,\mu) + c||^2 - \xi^2), \\ \qquad \text{otherwise}. \end{cases}$$

According to this definition, if the sequence $\{(\lambda_k, \mu_k)\}$ is close to the boundary $\partial\Omega_0$ from the interior of Ω_0, then when $k \to +\infty$, we have

$$(\lambda_k, \mu_k) \to \arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda, \mu) \in \partial\Omega_0, \tag{3.1}$$

or

$$\Psi(\lambda_k, \mu_k) \to -\infty, \text{ and } ||d(\lambda_k, \mu_k)|| \to \infty. \tag{3.2}$$

Further, if

$$\arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda, \mu) \in \partial\Omega_0, \tag{3.3}$$

then Hessian matrix $H(\lambda^*, \mu^*)$at the global solution d^* of the CDT subproblem is semi-positive definite or has one negative eigenvalue.

On the other hand, if $H(\lambda^*, \mu^*)$ has one negative eigenvalue at the global solution d^*, and the sequence $\{(\lambda_k, \mu_k)\}$ is close to the boundary $\partial\Omega_1$ from the interior of Ω_1, we will have

$$\Psi(\lambda_k, \mu_k) \to +\infty, \text{ and } ||d(\lambda_k, \mu_k)|| \to \infty, \text{ if } k \to +\infty \tag{3.4}$$

or the limit of the sequence will be a KKT point.

Lemma 3.1. *If $\mathcal{F}$ has strict interior points, then the set*

$$L_0(\lambda_0, \mu_0) = \{(\lambda, \mu) \,|\, \Psi(\lambda, \mu) \geq \Psi(\lambda_0, \mu_0), \quad (\lambda, \mu) \in \Omega_0\} \tag{3.5}$$

is bounded, where $(\lambda_0, \mu_0) \in \Omega_0$.

Proof. Because $\mathcal{F}$ has strict interior points, there exists $\bar{d}$ such that

$$||\bar{d}|| < \Delta, \quad \text{and} \quad ||A^T\bar{d} + c|| < \xi. \tag{3.6}$$

Under the assumption $(\lambda, \mu) \in \Omega_0$, $d(\lambda, \mu)$ is calculated to minimize $\Psi(\lambda, \mu)$, so we have

$$\Psi(\lambda, \mu) < \Phi(\bar{d}) + \frac{\lambda}{2}(||\bar{d}||^2 - \Delta^2) + \frac{\mu}{2}(||A^T\bar{d} + c||^2 - \xi^2). \tag{3.7}$$

Hence $\Psi(\lambda, \mu) \to -\infty$ if $\max\{\lambda, \mu\} \to +\infty$. Therefore the lemma is true.

Theorem 3.2. *For $(\lambda, \mu) \in \Omega_0 \cup \Omega_1$, assume that $d(\lambda, \mu)$ is defined by the equation (2.4). then $||d(\lambda, \mu)|| \to \infty$ only if (λ, μ) is very close to $\partial\Omega_0 \cup \partial\Omega_1$.*

Proof. There are two cases.

a). there does not exist $\lambda_s = c > 0$ which is the singular line(see [2]) of $H(\lambda, \mu)$.
In this case, we can find two points $(\lambda_a, 0)$ and $(0, \mu_b)$ such that $H(\lambda, \mu)$ is positive definite. We define

$$\Omega_\infty = \{(\lambda, \mu) \in \mathcal{R}^2_+ |\, \lambda > \lambda_a, \text{ or } \mu > \mu_b\} \tag{3.8}$$

Obviously, we only need prove that $||d(\lambda, \mu)||$ is bounded above in Ω_∞. For all $(\lambda, \mu) \in \Omega_\infty$, we have

$$||d(\lambda, \mu)|| \leq ||H(\lambda, \mu)^{-1}g|| + ||H(\lambda, \mu)^{-1}\mu Ac|| \tag{3.9}$$

the first term of the right-hand side of (3.9) is bounded above by

$$\max\{||H(\lambda_a, 0)^{-1}||, ||H(0, \mu_b)^{-1}||\}||g|| \tag{3.10}$$

On the other hand, if $\lambda > \lambda_a$, using the relation

$$Ac = AA^+Ac = AA^T(A^+)^Tc, \tag{3.11}$$

it can be shown that

$$||H(\lambda, \mu)^{-1}\mu Ac|| \leq ||H(\lambda_a, \mu)^{-1}\mu Ac|| \tag{3.12}$$
$$= ||(A^+)^Tc - H(\lambda_a, \mu)^{-1}H(\lambda_a, 0)(A^+)^Tc|| \tag{3.13}$$
$$\leq \big(1 + ||H(\lambda_a, 0)^{-1}||\, ||H(\lambda_a, 0)||\big)\, ||(A^+)^Tc|| \tag{3.14}$$

Similarly, we can show that

$$||H(\lambda, \mu)^{-1}\mu Ac|| \leq \big(1 + ||H(0, \mu_b)^{-1}||\, ||H(0, \mu_b)||\big)\, ||(A^+)^Tc|| \tag{3.15}$$

when $\mu > \mu_b$. Therefore $||d(\lambda, \mu)||$ is bounded above in Ω_∞.

b). There exists $\lambda = \lambda_s(\lambda_s > 0)$ which is the singular line of $H(\lambda, \mu)$.
According to the above proof, we only need consider the case that

$$\lambda < \lambda_s - \epsilon \text{ and } \mu \to \infty \tag{3.16}$$

where $\epsilon > 0$ is small enough. In this case, we can also show that (3.9) holds. It is trivial that

$$||H(\lambda, \mu)^{-1}g|| \leq \frac{||g||}{\epsilon} \tag{3.17}$$

about the second term of the right-hand side of (3.9), we have

$$\begin{aligned} ||H(\lambda, \mu)^{-1}\mu Ac|| &= ||(I - H(\lambda, \mu)^{-1}(B + \lambda I))(A^+)^Tc|| \\ &\leq (1 + (||B|| + \lambda_s)/\epsilon)\, ||(A^+)^Tc|| \end{aligned} \tag{3.18}$$

From (3.17)–(3.18), we have

$$||d(\lambda,\mu)|| \leq \frac{||g||}{\epsilon} + \left(1 + \frac{||B|| + \lambda_s}{\epsilon}\right) ||(A^+)^T c|| \tag{3.19}$$

This completes the proof of this lemma.

According to the definition of $\Psi(\lambda,\mu)$, we have

Corollary 3.3. $\nabla^2\Psi(\lambda,\mu)$ *is bounded above in the set* $L_0(\lambda_0,\mu_0)$*, where* $L_0(\lambda_0,\mu_0)$ *is defined by* (3.5)*.*

Proof. If (λ,μ) is very close to $\partial\Omega_0$ and $(\lambda,\mu) \in \Omega_0$, then from the definition of $\Psi(\lambda,\mu)$, we have $\Psi(\lambda,\mu) \to -\infty$ or $\max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda,\mu)$.

Further, at the first case, we have

$$(\lambda,\mu) \notin L_0(\lambda_0,\mu_0);$$

On the other hand, if it is in the latter case, we have $\nabla^2\Psi(\lambda,\mu)$ is bounded. Since $L_0(\lambda_0,\mu_0)$ is bounded and Theorem 3.2, we know that corollary is true.

Corollary 3.4. $\nabla^2\Psi(\lambda,\mu)$ *is bounded in the set*

$$L_1(\lambda_+,\mu_+) = \{(\lambda,\mu) \,|\, ||\nabla\Psi(\lambda,\mu)|| \leq ||\nabla\Psi(\lambda_+,\mu_+)||, (\lambda,\mu) \in \Omega_1\} \tag{3.20}$$

where

$$(\lambda_+,\mu_+) = arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda,\mu).$$

Corollary 3.4 can be shown from theorem 3.2 and the definition of dual function, so we omit the proof.

4. Algorithm

The algorithm presented below is iterative. At each iteration, an estimate of the Lagrange multiplier (or dual variable) (λ_k,μ_k) is known, then, an acceptable step $(\delta\lambda,\delta\mu))$ is calculated and the next iterate is $(\lambda_k+\delta\lambda,\mu_k+\delta\mu)$.

In detail, our algorithm consists of two stages. At the first stage, the goal is to maximize the dual function $\Psi(\lambda,\mu)$ in the set Ω_0, namely,

$$\max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda,\mu). \tag{4.1}$$

When this stage is over, we get

$$(\lambda_+,\mu_+) = \arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda,\mu).$$

If $d(\lambda_+,\mu_+)$ is the global solution d^* of the CDT subproblem, then algorithm will stop; Otherwise, we enlarge search region from Ω_0 to $\Omega_0 \cup \Omega_1$ and enter the second stage, now our goal is to find a feasible KKT point satisfying Lagrange multiplier $(\lambda^*,\mu^*) \in \Omega_1$. In one word, whether at the first stage or at the second stage, we try to search (λ,μ) in Ω_0 or Ω_1 such that the triple $(\lambda,\mu,d(\lambda,\mu))$ is a feasible KKT point of the CDT subproblem.

In addition, the successes of Gay[5] and Moré & Sorensen[8] on the trust region subproblem

$$\min g^T d + \tfrac{1}{2} d^T B d, \tag{4.2}$$

$$\text{s.t.} \qquad ||d|| \leq \Delta. \tag{4.3}$$

imply that the rational structure of $||d||$, $||A^T d + c||$ may be exploited.

So for the CDT subproblem (1.1)-(1.2), We apply Newton method to the equation system

$$w(\lambda,\mu) = \begin{pmatrix} 1/\Delta - 1/||d(\lambda,\mu)|| \\ 1/\xi - 1/||A^T d(\lambda,\mu) + c|| \end{pmatrix} = 0 \tag{4.4}$$

at each iteration, where $d(\lambda,\mu)$ is defined by

$$H(\lambda,\mu)d(\lambda,\mu) = -(g + \mu Ac). \tag{4.5}$$

That is to say, we solve

$$w(\lambda,\mu) + \nabla w(\lambda,\mu) \begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} = 0. \tag{4.6}$$

Direct calculation shows that

$$\nabla w(\lambda,\mu) = -\begin{pmatrix} dH^+d/||d||^3 & dH^+y/||d||^3 \\ dH^+y/||A^Td+c||^3 & yH^+y/||A^Td+c||^3 \end{pmatrix} \tag{4.7}$$

where $d = d(\lambda,\mu)$, $H = H(\lambda,\mu)$ and $y = y(\lambda,\mu)$.

Because the matrix $\nabla w(\lambda,\mu)$ may be singular, we use the generalized Newton step

$$\bar{\rho} = -\left(\nabla w(\lambda,\mu)\right)^+ w(\lambda,\mu) \tag{4.8}$$

On the other hand, when $\nabla w(\lambda,\mu)$ is singular, generalized Newton step $\bar{\rho}$ is only a minimizer of

$$||w(\lambda,\mu) + \nabla w(\lambda,\mu) \begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix}||$$

in the range space of $\nabla^2\Psi(\lambda,\mu)$; Further, when $\nabla^2\Psi(\lambda,\mu)$ is almost singular, generalized Newton step will be too large. Considering these situations, we define

$$Q(s) = s^T\nabla\Psi(\lambda,\mu) + \frac{1}{2}s^T\nabla^2\Psi(\lambda,\mu)s, \ s \in \mathcal{R}^2 \tag{4.9}$$

if either

$$\bar{\rho}^T\nabla\Psi(\lambda,\mu) < \frac{1}{M}||\nabla\Psi(\lambda,\mu)||^2 \quad \text{or} \quad ||\bar{\rho}|| > s \tag{4.10}$$

holds, we use a steepest ascent step

$$\hat{\rho} = \frac{1}{M}\nabla\Psi(\lambda,\mu) \tag{4.11}$$

where $M \geq |\theta_1| + |\theta_2|$ (θ_1, θ_2 are the diagonal elements of matrix $\nabla^2\Psi(\lambda,\mu)$), and s is a parameter updated at each iteration.

At the boundary, we search along the boundary if necessary. At the point $(0,\mu)^T$, if $w_1(\lambda,\mu) \geq 0$ (this may indicate that the case when the first constraint of the CDT subproblem is inactive) or if the first component of the calculated step ($\bar{\rho}$ or $\hat{\rho}$) is negative (such that the trial step is infeasible), we then use the "projected steepest ascent direction",

$$\begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} = \begin{pmatrix} 0 \\ -w_2(\lambda,\mu)/(\partial w_2(\lambda,\mu)/\partial\mu) \end{pmatrix}. \tag{4.12}$$

Similarly, at the boundary point $(\lambda,0)^T$ if $w_2(\lambda,\mu) \geq 0$ or if the second component of the trial step is negative, we use the step

$$\begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} = \begin{pmatrix} -w_1(\lambda,\mu)/(\partial w_1(\lambda,\mu)/\partial\lambda) \\ 0 \end{pmatrix}. \tag{4.13}$$

In the end, a step is truncated if it makes the new point $(\lambda+\delta\lambda, \mu+\delta\mu)$ infeasible. Namely, we choose the largest $t \in (0,1]$ that satisfies

$$(\lambda + t\delta\lambda, \mu + t\delta\mu) \in \mathcal{R}_+^2. \tag{4.14}$$

At the first stage, our condition for accepting a trial step is that

$$(\nabla\Psi(\lambda + t\delta\lambda, \mu + t\delta\mu))^T \begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} \geq 0, \tag{4.15}$$

$$\Psi(\lambda + t\delta\lambda, \mu + t\delta\mu) \geq \Psi(\lambda, \mu) + vt(\nabla\Psi(\lambda, \mu))^T \begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} \tag{4.16}$$

hold, where $v \in (0, 0.5)$ is a preset constant.

If a trail step $\bar{\rho}$ is unacceptable, we replace it by $\hat{\rho}$. Then we keep increasing M by twice until the step $\hat{\rho}$ is acceptable. It can be seen that if

$$M \geq \max_{(\lambda,\mu)\in L_0(\lambda_0,\mu_0)} ||\nabla^2\Psi(\lambda, \mu)||, \tag{4.17}$$

(4.15) will hold.

On the other hand, if (4.15) holds and (4.16) does not hold, then we can use the same method as Yuan[13] and show

$$\Psi(\lambda + t\delta\lambda, \mu + t\delta\mu) \geq \Psi(\lambda, \mu) + \frac{2v(1-v)}{M} || \begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} ||^{-2} \left[(\nabla\Psi(\lambda, \mu))^T \begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} \right]^2 \tag{4.18}$$

holds, where M satisfies (4.17). That is to say, either (4.16) or (4.18) holds, we accept trial step.

Similarly, at the first stage, if trial step (4.12) is unacceptable, we choose

$$\delta\mu := \delta\mu \frac{w_2(0, \mu)}{w_2(0, \mu + \delta\mu) - w_1(0, \mu)}. \tag{4.19}$$

if (4.13) is unacceptable, we set

$$\delta\lambda := \delta\lambda \frac{w_1(\lambda, 0)}{w_1(\lambda + \delta\lambda, 0) - w_1(\lambda, 0)}. \tag{4.20}$$

At the second stage, we use

$$F(\lambda, \mu) = \frac{1}{2}||r(\lambda, \mu)||^2, \tag{4.21}$$

as merit function, and the goal is to minimize $F(\lambda, \mu)$, where

$$r(\lambda, \mu) = \begin{cases} e_2 e_2^T \nabla\Psi(\lambda, \mu) & \text{if } \lambda = 0 \text{ and } w_1(\lambda, \mu) \leq 0, \\ e_1 e_1^T \nabla\Psi(\lambda, \mu) & \text{if } \mu = 0 \text{ and } w_2(\lambda, \mu) \leq 0, \\ (0,0)^T & \text{if } \lambda = \mu = 0, w_1(\lambda, \mu) \leq 0 \text{ and } w_2(\lambda, \mu) \leq 0, \\ \nabla\Psi(\lambda, \mu) & \text{otherwise.} \end{cases} \tag{4.22}$$

It is obvious that $F(\lambda, \mu)$ is piecewise continuous function. Now our condition for accepting trial step is

$$F(\lambda_k + \sigma^l t\delta\lambda, \mu_k + \sigma^l t\delta\mu) \leq F(\lambda_k, \mu_k) + b_1\sigma^l t(J_k^T r(\lambda_k, \mu_k))^T \begin{pmatrix} \delta\lambda \\ \delta\mu \end{pmatrix} \tag{4.23}$$

where $J_k = J(\lambda_k, \mu_k)$ is Jacobi matrix, $\sigma = 1/2$ and l is the least nonnegative integer satisfying (4.23).

From Corollary 3.4, we know $J(\lambda, \mu)$ is Lipschitz continuous in $L_1(\lambda_+, \mu_+)$. Hence according to the analysis of Yuan[11], we have

$$\sigma^l \geq c||J_k^T r(\lambda_k, \mu_k)||^2 cos\theta_k \tag{4.24}$$

holds, where θ_k is the angle between $J_k^T F(\lambda_k, \mu_k)$ and trial step $t(\delta\lambda, \delta\mu)$.

In our algorithm, we use variable *solve* to divide the first stage from the second stage. When $solve = .false.$, the algorithm is in the first stage; otherwise if $solve = .true.$, the algorithm enters the second stage.

In order to judge the eigenvalues of H_k and solve the equation (4.5), we factorize H_k with pivoting [6] at each iteration,

$$H(\lambda_k, \mu_k) = R_k^T \Lambda_k R_k. \tag{4.25}$$

where Λ_k is a diagonal matrix and all the diagonal elements are 1 or -1. That is, if one diagonal element of H_k is equal to -1, then

$$(\lambda_k, \mu_k) \in \Omega_1. \tag{4.26}$$

Now the detail of the algorithm can be given as follows.

Algorithm 4.1.

Step 0 *Given* $\lambda_1 = ||B||_1,\ \mu_1 = ||B||_1,\ solve = .false., \varepsilon > 0,$ *and* $k = 1$.

Step 1 *Factorize* $H(\lambda_k, \mu_k) = R_k^T \Lambda_k R_k$ *with pivoting.*

step 2 *Calculate* d_k *from (4.5).*

step 3 *Calculate the least singular value* σ_{min} *and the corresponding vector* $\tilde{u}(||\tilde{u}|| = 1)$ *of* R_k^T. *if* $\sigma_{min} < \varepsilon$, *calculate* $\tilde{\tau}$ *such that*

$$\lambda_k(||d_k + \tilde{\tau}\tilde{u}||^2 - \Delta^2) + \mu_k(||A^T(d_k + \tilde{\tau}\tilde{u}) + c||^2 - \xi^2) = 0. \tag{4.27}$$

If convergence, then stop; If

$$(||d_k + \tilde{\tau}_+\tilde{u}|| - \Delta)(||d_k + \tilde{\tau}_-\,\tilde{u}|| - \Delta) < 0, \tag{4.28}$$

then $solve = .true.$

step 4 *Calculate*

$$w_1^{(k)} = \frac{1}{\Delta} - \frac{1}{||d_k||_2}, \qquad w_2^{(k)} = \frac{1}{\xi} - \frac{1}{||A^T d_k + c||_2};$$

if convergence, then stop;

$$M_k = \max\{M_{k-1}, |d_k^T H_k^+ d_k| + |y_k^T H_k^+ y_k|\}; \qquad s_k = \max\{s_{k-1}, ||w^{(k)}||/M_k\}.$$

if $\lambda_k = 0$ *and* $w_1^{(k)} \leq 0$, *go to step 6; if* $\mu_k = 0$ *and* $w_2^{(k)} \leq 0$, *go to step 7.*

step 5 *Calculate the generalized Newton step* $\bar{\rho}$.
If (4.10) holds, set $(\delta\lambda, \delta\mu)^T = \hat{\rho}$; *otherwise, set* $(\delta\lambda, \delta\mu)^T = \bar{p}$.
If $\lambda_k = 0$ *and* $\delta\lambda < 0$, *go to step 6; If* $\mu_k = 0$ *and* $\delta\mu < 0$, *go to step 7.*

step 6 $\delta\lambda = 0,\ \delta\mu = -w_2^{(k)}/\partial w_2^{(k)}$, *go to step 8.*

step 7 $\delta\lambda = -w_1^{(k)}/\partial w_1^{(k)},\ \delta\mu = 0.$

step 8 *Truncate the step if necessary(calculate* $t_k = 1/\max\{1, -\delta\lambda/\lambda_k, -\delta\mu/\mu_k\}$*; and* $(\delta\lambda, \delta\mu)^T := t_k(\delta\lambda, \delta\mu)^T$ *).*

step 9 *Carry out line search. calculate the smallest no-negative integer* $I(k)$ *such that accepted criteria hold for the trial step* $2^{-I(k)}(\delta\lambda, \delta\mu)^T$*; Set* $(\delta\lambda, \delta\mu)^T = 2^{-I(k)}(\delta\lambda, \delta\mu)^T$ *and* $M_k = 2^{I(k)} M_k / t_k$ *if* $I(k) > 0$*.*

Step 10 *Factorize* $H(\lambda_k + \delta\lambda, \mu_k + \delta\mu) = R_{k+1}^T \Lambda_{k+1} R_{k+1}$ $(\Lambda_{k+1} =$ *diag* $(1, \cdots, 1, \pm 1))$*with pivoting, and the last diagonal element of* Λ_{k+1} *is equal to* -1 *only if solve = .true. holds. If* $H \geq 0$ *or* H *has one negative eigenvalue but solve = .true., go to step 11; otherwise* $(\delta\lambda, \delta\mu)^T = (\delta\lambda, \delta\mu)^T/2$*, go to step 9.*

step 11 *Set* $\lambda_{k+1} = \lambda_k + \delta\lambda,\ \mu_{k+1} = \mu_k + \delta\mu$*; Set* $k = k + 1$ *and go to step 2.*

At **step 4** of Algorithm 4.1, the stopping criterion is

$$r(\lambda_k, \mu_k) \geq -\epsilon\, e \quad \text{and} \quad (\lambda_k r_1(\lambda_k, \mu_k), \mu_k r_2(\lambda_k, \mu_k))^T \leq \epsilon\, e, \tag{4.29}$$

where $e = (1, 1)^T$; At **step 3**, when

$$\tilde{\tau}^2 \tilde{u}^T H_k \tilde{u} \leq \sigma \left| d_k^T H_k d_k + \lambda_k \Delta^2 + \mu_k(\xi^2 - ||c||^2) \right| \quad \text{and} \quad |\, ||d_k + \tilde{\tau}\tilde{u}|| - \Delta| \leq \epsilon\Delta \tag{4.30}$$

hold, Algorithm 4.1 will halt.

Lemma 4.2. *Let* $0 < \sigma < 1$, $\varepsilon \geq 0$ *be given and suppose that*

$$H(\lambda, \mu) = B + \lambda I + \mu A A^T = R^T R, \tag{4.31}$$

$$H(\lambda, \mu) d = -(g + \mu A c),\ \lambda \geq 0,\ \mu \geq 0 \tag{4.32}$$

Let $\tilde{u} \in R^n(||\tilde{u}|| = 1)$ *and* $\tilde{\tau}$ *satisfy*

$$\lambda\left(||d + \tilde{\tau}\tilde{u}|| - \Delta\right) + \mu\left(||A^T(d + \tilde{\tau}\tilde{u}) + c)|| - \xi\right) = 0 \tag{4.33}$$

$$\tilde{\tau}^2 \tilde{u}^T H(\lambda, \mu) \tilde{u} \leq \sigma \left| d^T H(\lambda, \mu) d + \lambda\Delta^2 + \mu(\xi^2 - ||c||^2) \right| \tag{4.34}$$

$$|\, ||d + \tilde{\tau}\tilde{u}|| - \Delta| \leq \varepsilon\Delta \tag{4.35}$$

Then

$$|\, \Phi^* - \Phi(d + \tilde{\tau}\tilde{u})\, | \leq \sigma\, |\Phi^*| \tag{4.36}$$

where Φ^* *is the optimal value of (1.1)–(1.2).*

Proof. Let $M = \frac{1}{2}\left(d^T H(\lambda, \mu) d + \lambda\Delta^2 + \mu(\xi^2 - ||c||^2)\right)$. First, note that for any $\tilde{u} \in R^n$,

$$\Phi(d + \tilde{\tau}\tilde{u}) = -\frac{1}{2}\left[d^T H d + \lambda ||d + \tilde{\tau}\tilde{u}||^2 + \mu(||A^T(d + \tilde{\tau}\tilde{u}) + c)||^2 - ||c||^2)\right] + \frac{1}{2}\tilde{\tau}^2 \tilde{u}^T H \tilde{u} \tag{4.37}$$

where $H = H(\lambda, \mu)$. Then for any $\tilde{\tau}, \tilde{u}$ satisfying (4.33)–(4.35),we have

$$-\Phi(d + \tilde{\tau}\tilde{u}) \geq M - \sigma\, |\, M\, | \tag{4.38}$$

Furthermore, if $\Phi^* = \Phi(d + \tilde{\tau}^* \tilde{u}^*)$ where $d + \tilde{\tau}^* \tilde{u}^* \in \mathcal{F}$, then (4.37) implies

$$-\Phi(d + \tilde{\tau}^* \tilde{u}^*) \leq M. \tag{4.39}$$

The last two inequalities yield this Lemma.

Lemma 4.2 shows that $d + \tilde{\tau}\tilde{u}$ is a nearly optimal solution to problem (1.1)–(1.2).

5. Convergence Properties

From the definition of $r(\lambda, \mu)$, we also use the notation

$$z_k = ||r(\lambda_k, \mu_k)||, \tag{5.1}$$

for all k. It can be shown that Algorithm 4.1 converges if and only if

$$z_k \to 0 (k \to \infty). \tag{5.2}$$

First, we consider the angle between trial step and gradient direction.

Lemma 5.1. *Let $\varrho_k = (\delta\lambda, \delta\mu)^T$ be the trial step at the k-th iteration, then we have*

$$\alpha_k = \cos(\varrho_k, \nabla\Psi_k) \geq \min\{1, z_k/M_k s_k\}, \tag{5.3}$$

where $\nabla\Psi_k = \nabla\Psi(\lambda_k, \mu_k)$.

Proof. From (5.1), we have

$$||\nabla\Psi_k|| \geq z_k, \tag{5.4}$$

If $\varrho_k = \hat{\rho}$, then $\alpha_k = 1$, (5.3) holds.

When $\varrho_k = \bar{\rho}$, according to (4.10), we get

$$\alpha_k = \frac{\bar{\rho}^T \nabla\Psi_k}{||\bar{\rho}|| \cdot ||\nabla\Psi_k||} \geq \frac{||\nabla\Psi_k||}{M_k||\bar{\rho}||} \geq \frac{z_k}{M_k s_k}, \tag{5.5}$$

namely, (5.3) also holds.

Now we consider the case that ϱ_k is defined by (4.12). There are two cases.

1) $w_1(\lambda_k, \mu_k) \leq 0$. It can be shown that

$$\alpha_k = |e_2^T \nabla\Psi_k|/||\nabla\Psi_k|| \geq z_k/||\nabla\Psi_k|| \geq z_k/M_k s_k, \tag{5.6}$$

2) $w_1(\lambda_k, \mu_k) > 0$, $e_1^T \bar{\rho} < 0$. in this case, since

$$(e_2^T \nabla\Psi_k) e_2^T \bar{\rho} = \bar{\rho}^T \nabla\Psi_k - (e_1^T \nabla\Psi_k) e_1^T \bar{\rho} \geq \bar{\rho}^T \nabla\Psi_k \geq ||\nabla\Psi_k||^2/M_k, \tag{5.7}$$

it follows that

$$\alpha_k = |e_2^T \nabla\Psi_k|/||\nabla\Psi_k|| \geq ||\nabla\Psi_k||/(||\bar{\rho}||M_k) \geq z_k/(M_k s_k). \tag{5.8}$$

In the same method, we can show that (5.3) also holds when trial step is decided by (4.13). This completes the proof of theorem.

The following lemma shows that $I(k) = 0$ for sufficiently large k at the first stage.

Lemma 5.2. *At the first stage, there are only finitely many k such that $I(k) > 0$. Consequently, M_k is bounded above.*

Proof. Firstly, from Theorem 3.1 and Theorem 3.2, $\nabla^2\Psi(\lambda, \mu)$ is bounded in the set $L_0(\lambda_1, \mu_1)$, where (λ_1, μ_1) is the initial point of Algorithm 4.1. Namely, we have

$$||\nabla^2\Psi(\lambda, \mu)|| \leq K, \quad \forall(\lambda, \mu) \in L_0(\lambda_1, \mu_1). \tag{5.9}$$

If conclusion of Lemma does not hold, that is to say, there are infinitely many k such that $I(k) > 0$ at the first stage, we can show

$$\lim_{k\to\infty} M_k = \infty. \tag{5.10}$$

Hence there exists an integer k_0 such that

$$M_k \geq M, \quad \forall\, k \geq k_0. \tag{5.11}$$

where M is the same as equation (4.17). Therefore, at the first stage, $\hat{\rho}$ is always an acceptable step for $k \geq k_0$. which contradicts (5.10). Thus the lemma is true.

Using the above results, we can prove theorem below.

Theorem 5.3. *The sequence* $\{(\lambda_k, \mu_k)^T; k = 1, 2, 3 \cdots\}$ *was generated by Algorithm 4.1 at the first stage, then* $\{(\lambda_k, \mu_k)^T\}$ *converges to*

$$arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda, \mu). \tag{5.12}$$

Proof. At the first stage, if $(\lambda_k, \mu_k) \in \text{int}\Omega_0$ and $(\lambda_k + t\delta\lambda, \mu_k + t\delta\mu)$ generated by Algorithm 4.1 is close to or goes beyond the boundary $\partial\Omega_0$, then unique possibility is

$$\arg \max_{(\lambda,\mu)\in\Omega_0} \Psi(\lambda, \mu) \in \partial\Omega_0. \tag{5.13}$$

In this case, we must shorten the trial step through **step 11** of Algorithm 4.1, and consequently the sequence $\{(\lambda_k, \mu_k)\}$ will converge to the maximizer on the boundary $\partial\Omega_0$. Therefore the first stage of Algorithm 4.1 will be over. That is to say, Shortening trial step by **step 11** of Algorithm 4.1 will not occur infinitely.

Due to (4.15)-(4.16), we deduce

$$\Psi(\lambda_{k+1}, \mu_{k+1}) > \Psi(\lambda_k, \mu_k). \tag{5.14}$$

further, we have that $\lim_{k\to\infty} z_k = 0$ and $\lim_{k\to\infty} \inf z_k = 0$ are equivalent.
Assume that theorem is false, that is, there exists $\tau > 0$ such that

$$z_k \geq \tau \tag{5.15}$$

for all k. In this case, we know that contracting trial step by **step 11** will not occur.

Now we consider the lower boundary of $||\varrho_k||$. When $\varrho_k = \hat{\rho}$, it is obvious from (5.4) that

$$||\varrho_k|| = ||\nabla\Psi_k||/M_k \geq z_k/M_k; \tag{5.16}$$

if $\varrho_k = \bar{\rho}$, we deduce from Assumption 2.3 that

$$||\varrho_k|| \geq c_2 z_k/||\nabla^2\Psi_k|| \tag{5.17}$$

where $c_2 = 2 \cdot \min\{\frac{c_0^2}{\xi(\xi+c_0)}, \frac{c_0^2}{\Delta(\Delta+c_0)}\}$.

If ϱ_k is defined by (4.12), there are two cases.

i) $w_1(\lambda_k, \mu_k) \leq 0$, so it is easy to see that

$$||\varrho_k|| \geq c_2 z_k/y_k^T H_k^+ y_k \geq c_2 z_k/M_k. \tag{5.18}$$

ii) $w_1(\lambda_k, \mu_k) > 0$ and $e_1^T\bar{\rho} < 0$ hold, we have

$$\begin{aligned} e_2^T w(\lambda_k, \mu_k) e_2^T\bar{\rho} &= \bar{\rho}^T w(\lambda_k, \mu_k) - e_1^T w(\lambda_k, \mu_k) e_1^T\bar{\rho} \\ &\geq \bar{\rho}^T w(\lambda_k, \mu_k) \geq c_2 c_3 z_k^2/M_k. \end{aligned} \tag{5.19}$$

where

$$c_3 = 1/\max_{(\lambda,\mu)\in L_0(\lambda_0,\mu_0)}\{\Delta||d(\lambda,\mu)||(\Delta+||d(\lambda,\mu)||), \xi||A^T d(\lambda,\mu)+c||(\xi+||A^T d(\lambda,\mu)+c||)\},$$

and from Theorem 3.2, c_3 is a constant. therefore we deduce that

$$||\varrho_k|| = |e_2^T w(\lambda_k,\mu_k)|/y_k^T H_k^+ y_k \geq c_2 c_3 \frac{z_k^2}{M_k^2 s_k} \tag{5.20}$$

That is to say, if trial step is defined by (4.12), then we can show that

$$||\varrho_k|| \geq \min\{1, c_3\cdot z_k/(M_k s_k)\} c_2 z_k/M_k. \tag{5.21}$$

Furthermore if trial step defined by (4.12) was modified by (4.19), we can get

$$||\varrho_k|| \geq \min\{1, c_3\cdot z_k/(M_k s_k)\} c_2 c_3 z_k^2/(2\bar{M} M_k), \tag{5.22}$$

where $\bar{M} = \max_{(0,\mu)\in\Omega_0}|w_2(0,\mu)|$.

Using similar method, when trial step was derived from (4.13), the acceptable trial step is such that

$$||\varrho_k|| \geq \min\{1, c_3\cdot z_k/(M_k s_k)\} c_2 c_3 z_k^2/(2\hat{M} M_k), \tag{5.23}$$

where $\hat{M} = \max_{(\lambda,0)\in\Omega_0}|w_1(\lambda,0)|$.

According to (5.15)-(5.23) and Lemma5.2 , there is $\beta > 0$ such that

$$||\varrho_k|| \geq \beta \tag{5.24}$$

for all k.

From (4.15)-(4.16), we deduce that

$$\begin{gathered}\Psi(\lambda_{k+1},\mu_{k+1}) - \Psi(\lambda_k,\mu_k) \geq \\ \min\{vt_k\beta z_k \min\{1, z_k/M_k s_k\}, \tfrac{2v(1-v){z_k}^2}{M^*}[\min\{1, z_k/M_k s_k\}]^2\}\end{gathered} \tag{5.25}$$

where $M^* = \max_{(\lambda,\mu)\in L_0(\lambda_1,\mu_1)}||\nabla^2\Psi(\lambda,\mu)||$.

Because $\Psi(\lambda,\mu)$ in the set Ω_0 is bounded above, we get

$$\sum_{k=1}^{\infty} t_k < \infty. \tag{5.26}$$

which implies that the sequence $\{(\lambda_k,\mu_k)^T\}$ converges, and

$$\lim_{k\to+\infty}(\lambda_k,\mu_k)^T = (0,0)^T. \tag{5.27}$$

Namely, B is positive definite and two constraints are inactive at the global solution d^* for CDT subproblem. This contradicts Assumption 2.3, so theorem is true.

Due to Chen & Yuan[2], we know, Algorithm 4.1 either finds the global solution d^* for CDT subproblem (1.1)-(1.2) at the first stage or enters into the second stage with the condition of (4.28).

If there is the global solution d^* with positive semi-definite $H(\lambda^*,\mu^*)$ for CDT subproblem (1.1)-(1.2), then Algorithm will stop at the first stage. Namely, if Algorithm 4.1 enters into the second stage, it implies that there is no positive semi-definite $H(\lambda^*,\mu^*)$ at the global solution d^*; At this time, we can conclude that $H(\lambda^*,\mu^*)$ has one negative eigenvalue and the corresponding vectors d^* and y^* are linearly independent.

The following theorem shows that Algorithm 4.1 will stop at the second stage.

Theorem 5.4. *Suppose that the sequence $\{(\lambda_k, \mu_k)^T; k = 1, 2, 3\cdots\}$ is generated by Algorithm 4.1 at the second stage, then either Algorithm will stop after finite iteration or $\{(\lambda_k, \mu_k)\}$ will converges in the sense that*

$$\lim_{k\to\infty} ||J_k^T r_k|| = 0. \tag{5.28}$$

Proof. From the above analysis, when the second stage of algorithm begins, d^* and y^* are linearly independent at all the KKT points where H^* has one negative eigenvalue. Hence $\nabla^2\Psi(\lambda, \mu)$ is nonsingular in the set $\text{int}\Omega_1$.

If $\{(\lambda_k, \mu_k)\} \in \text{int}\Omega_1$, and $\{(\lambda_k + t\delta\lambda, \mu_k + t\delta\mu)\}$ is sufficiently close to $\partial\Omega_1$, then the only impossibility is that there is a KKT point of the CDT subproblem in the set $\partial\Omega_1$; In this case, algorithm will satisfy the stopping criterion after finite iterations. Otherwise, we suppose that $\partial\Omega_1$ has no KKT point of the CDT subproblem. $\{(\lambda_k, \mu_k)\}$ generated by Algorithm 4.1 will be not close to $\partial\Omega_1$. without loss of generality, we can assume that

$$|\lambda_{min}(H(\lambda, \mu))| > \epsilon, \tag{5.29}$$

for all $(\lambda, \mu) \in L_1(\lambda_+, \mu_+)$, where $|\lambda_{min}(H(\lambda, \mu))|$ is the absolute minimizer of the eigenvalues of $H(\lambda_k, \mu_k)$.

If theorem is false, we have

$$||J_k^T r_k|| > \tau. \tag{5.30}$$

where $\tau > 0$ is a constant.

According to corollary 3.4 and (5.29), we have

$$m_0 \le ||J_k|| \le M_0. \tag{5.31}$$

Hence (5.30) shows also that

$$||r_k|| \ge \tau_1. \tag{5.32}$$

From (4.23) and (4.24), we have

$$F_{k+1} \le F_k - b_2(\cos\theta_k)^2, \tag{5.33}$$

where $b_2 > 0$ is a constant.

According to Assumption 2.3, (5.29)-(5.32), using the similar method as Lemma 5.1, we can show that $\cos\theta_k$ has lower bound. So

$$F_k \to -\infty,$$

the contradiction occurs. Therefore theorem is true.

The local convergence result is as follows:

Theorem 5.5. *Suppose that $\{(\lambda_k, \mu_k)^T; k = 1, 2, 3\cdots\}$ generated by Algorithm 4.1 converges to $(\lambda^*, \mu^*)^T$, further, if $(\lambda^*, \mu^*)^T$ such that $\lambda^* + \mu^* > 0$, then $(\lambda_k, \mu_k)^T$ converges to $(\lambda^*, \mu^*)^T$ Q-superlinearly.*

Proof. At k-th iteration, Newton step $\bar{\rho}$ can be written as

$$\bar{\rho} = -(\nabla^2\Psi(\lambda_k, \mu_k))^+ \begin{pmatrix} \frac{||d_k||^2}{\Delta(\Delta+||d_k||)}(||d_k||^2 - \Delta^2) \\ \frac{||A^T d_k + c||^2}{\xi(\xi+||A^T d_k + c||)}(||A^T d_k + c||^2 - \xi^2) \end{pmatrix} \tag{5.34}$$

It is easy to see that

$$\bar{\rho} \to -(\nabla^2\Psi(\lambda_k, \mu_k))^+ \nabla\Psi(\lambda_k, \mu_k).$$

Hence if two constraints are active at the point (λ^*, μ^*), then Q-superlinear convergence follows. Now we consider the case $\lambda^* = 0, ||d^*||^2 - \Delta^2 < 0$. In this case, trial step will be defined by (4.12), and this trial step is the Newton-Raphson step of the equation

$$\frac{1}{\xi} - \frac{1}{||A^T d(0,\mu) + c||^2} = 0$$

so Q-suplinear convergence follows. Similarly, when $\mu^* = 0, ||A^T d^* + c||^2 - \xi^2 < 0$, trial step defined by (4.13) will also be Q-suplinear. This completes the proof of theorem.

6. The description of examples

About the CDT subproblem, there are no standard test examples. Hence we list the examples used by this paper as follows.

Problem 1–5. denote Yuan's problems(see [13]),where B is positive definite.

Problem 6. $\xi = 2,\ \Delta = \frac{\sqrt{205}}{12},\ B = \text{diag}[-3,-2,-1,0],\ A = (I_{3\times 3}\ 0_{1\times 3})^T,\ c = (0, 1/2, 1/3)^T,$ $g = (-1, -1/2, -1/3, -1/4)^T$.

Problem 7–9. just replace B with $B - 2I$ or $B - I$ in Yuan's problem $3 - -5$ such that B is indefinite.

Problem 10. $\xi = \sqrt{5},\ \Delta = \sqrt{30},\ B = \text{diag}[-1,-2,-3,-4],\ A = (I_{2\times 2}\ \ 0_{2\times 2})^T,\ c = (-2, 0)^T,\ g = (-2, -6, -3, 0)^T$

Problem 11. $\xi = \sqrt{5},\ \Delta = 1,\ B = \text{diag}[-1,-2],\ A = I,\ c = (1,1)^T,\ g = (-2,-1)^T$.

Problem 13. only changes A into $2I$ in problem 6.

Problem 14. $B = \text{diag}[-50,-2], A = \text{diag}[5, 1/5], g = (-10,-1)^T, c = (1, -2/5)^T, \xi = 1, \Delta = 1$.

Problem 12,15,16. denote example 7.1, 7.3 and an extension of example 7.3 respectively in [4].

Problem 17. just substitutes $g = (24, -27)^T$ for "g" of example 7.1 in [4].

Remark. In problem $1 - 5$, B is positive definite; In problem $15 - 17$, H has one negative eigenvalue at the solution. In other problems , B is indefinite but H is positive semi-definite.

7. Numerical Results

Our algorithm is implemented in Fortran77, and test results are presented in the following three tables. To see the behavior of the algorithm, we include in the table the two quantities "NT/FC", where "NT" is the number of iteration needed for algorithm to reach a solution for each problem, and "FC" denotes the number of Cholesky factorization. Listed in the table are also the dual variables and primal solutions. Moreover, in **Algorithm 4.1**, the parameter ε controlling precision in calculation is chosen as 10^{-6}.

In the table 7.1, we give numerical results of examples form Yuan[13]. For these examples, B is positive definite. For example 2 and 3, only one constraint is active. In the end, our algorithm arrived at the global solutions.

Table 7.1 $B > 0$

np	n/m	NT	FC	dual variable	primal solution
1	4-1	2	2	0.1001607266D+01 0.3000321453D+01	0.4998714683D+00,-0.4995985060D+00 -0.4995985060D+00,-0.4995985060D+00
2	4-2	5	5	0.0000000000D+00 0.4296232629D+00	0.6010300392D+00, 0.2057932222D+00 0.0000000000D+00, 0.0000000000D+00
3	4-3	4	4	0.6071188494D+01 0.0000000000D+00	0.4242568279D+00, 0.5656757706D+00 0.7070947132D+00, 0.0000000000D+00
4	4-4	4	4	0.2633710837D+01 0.8301411093D+00	0.8525471460D+00, -0.2878664443D+00 -0.3040366588D+00, -0.3128226952D+00
5	4-4	4	4	0.1161211825D+01 0.2289727569D+01	0.5826944230D+00, -0.4721314020D+00 -0.4956259653D+00,0.4382171756D+00

There are nine examples in table 7.2, where B is indefinite but $H(\lambda^*, \mu^*)$ is positive semi-definite at the global solution d^*. Finally our algorithm succeed in finding their global solution. For example 6,7 and 13, only one constraint is active at the global solution.

Table 7.2 B is indefinite but $H(\lambda^*, \mu^*) \geq 0$

np	n/m	NT	FC	dual variable	primal solution
6	4-3	4	4	0.3865762605D+01 0.0000000000D+00	0.1155051044D+01,0.2679869340D+00 0.1163157523D+00,0.6467029292D-01
7	4-3	5	5	0.8071773746D+01 0.0000000000D+00	0.4242217169D+00,0.5656289558D+00 0.7070361948D+00,0.0000000000D+00
8	4-4	5	5	0.3633711119D+01 0.8301591742D+00	0.8525449044D+00,-0.2878689410D+00 -0.3040392942D+00,-0.3128254059D+00
9	4-4	5	5	0.2161452708D+01 0.2289573110D+01	0.5827110864D+00,-0.4720776368D+00 -0.4955687127D+00,0.4381789413D+00
10	4-2	5	9	0.4000000000D+01 0.1262561988D+01	0.1061597224D+01,0.1839045518D+01 0.3000000000D+01,0.4089455756D+01
11	2-2	3	5	0.1001027412D+01 0.9999180592D+00	0.9991372850D+00,0.8666655968D-01
12	2-2	5	5	0.5076923137D+01 0.1800000527D+01	0.1561981421D-05,-0.1299999980D+02
13	2-2	5	5	0.0000000000D+00 0.5854860215D+00	0.6177812964D+00,-0.5000000000D+00
14	2-2	9	15	0.2080070959D+01 0.2001715767D+01	-0.4040967637D-02, 0.9999854633D+00

As to three examples in table 7.3, B is indefinite and $H(\lambda^*, \mu^*)$ has one negative eigenvalue at the global solution d^*. In these three examples, our algorithm also finds the global solution, but theoretically the algorithm only guarantee to reach some local solution in this situation.

Table 7.3 B and $H(\lambda^*, \mu^*)$ have negative eigenvalues

np	n/m	NT	FC	dual variable	primal solution
15	2-2	8	10	0.0000000000D+00 0.3000005594D+01	-0.9999972031D+00, -0.9999888123D+00
16	3-3	7	9	0.0000000000D+00 0.3002702059D+01	-0.9986507936D+00, -0.9945812412D+00 0.0000000000D+00
17	2-2	10	16	0.9223248254D+00 0.1800890330D+01	0.2326383307D-02, -0.1300465547D+02

Acknowledgment. The authors would like to thank anonymous referees for their helpful comments.

References

[1] M.R. Celis, J.E. Dennis and R.A. Tapia, A trust region algorithm for nonlinear equality constrained optimization, In: P.T.Boggs, R.H.Byrd and R.B.Schnable, eds., *Numerical Optimization, SIAM*, Philadephia, 1985, 71-82.

[2] X.D. Chen and Y. Yuan, On maxima of dual function of the CDT subproblem, *J. Comp. Math.,* **19** (2001), 113-124.

[3] X.D. Chen and Y. Yuan, Optimality conditions for CDT subproblem, In: Y.Yuan, ed., Numerical Linear Algebra and Optimization, Science Press, Beijing, New York, 111-121.

[4] X.D. Chen and Y. Yuan, On local solutions of CDT subproblem, *SIAM J. Opt.,* **10** (1999), 359-383.

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

[6] G.H. Golub, C.F. Van Loan eds., Y. Yuan, ect. trans, Matrix Computation, Science Press, Beijing 2001.

[7] Gai-di Li, About KKT points of the CDT subproblem, submitted.

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

[9] J.M. Peng and Y. Yuan, Optimality conditions for the minimization of a quadratic with two quadratic constraints, *SIAM, J.Opti.,* **7** (1997), 579-594.

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

[11] Y. Yuan, Problems on convergence of unconstrained optimization algorithms, In:Y. Yuan, ed., *Numerical Linear Algebra and Optimization,* Science Press, Beijing, New York, 1999, 95-107.

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

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

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