Matrix computation problems in trust region algorithms for optimization *

Ya-xiang Yuan

State Key Laboratory of Scientific and Engineering Computing

Institute of Computational Mathematics and Scientific/Engineering Computing

Chinese Academy of Sciences, POB 2719, Beijing 100080, China

Abstract

Trust region algorithms are a class of recently developed algorithms for solving optimization problems. The subproblems appeared in trust region algorithms are usually minimizing a quadratic function subject to one or two quadratic constraints. In this paper we review some of the widely used trust region subproblems and some matrix computation problems related to these trust region subproblems.

Key words: optimization, trust region subproblem, matrix computation.

1. Introduction

Trust region algorithms are a class of recently developed algorithms for solving optimization problems. At each iteration of a trust region algorithm, a trial step is computed by solving a trust region subproblem, which is normally an approximation to the original optimization problem with a trust region constraint which prevents the trial step being too large. Usually, the trust region constraint has the form:

$$\|d\| \leq \Delta \tag{1.1}$$

where $\Delta > 0$ is the trust region bound.

For unconstrained optimization, the subproblems appeared in trust region algorithms are usually to minimize a quadratic function which is a quadratic

*Research partially supported by Chinese NSF grants 19525101, 19731010 and State Key project 96-221-04-02-02.

approximation to the objective function subject to the trust region constraint (1.1).

For constrained optimization. There are mainly three different types of trust region subproblems. The first type is a null space subproblem, where a quadratic model function is minimized in the null space of the linearized constraints subject to the trust region constraint (1.1) (for example, see [1] and [16]). The null space subproblem is basically the same as the standard trust region subproblem for unconstrained optimization. The second type of subproblems of trust region algorithms for equality constrained optimization is the so called CDT subproblem, which minimizes the quadratic model function subject to the trust region bound condition (1.1) and an additional quadratic constraint which has the form:

$$||A^T d + c|| \leq \xi \tag{1.2}$$

where $\xi \geq 0$ is a parameter, c is the constraint values at the current iterate point, and A is the gradient matrix of the constraints. The constraint condition (1.2) forces the sum of squares of the linearized constraint violations to reduce. The third type of trust region subproblems are exact penalty function type subproblems. Such a subproblem seeks a minimizer of the sum of a nonsmooth piece-wise linear function and a quadratic function within the trust region (1.1) (for example, see [12], [13] and [26]).

Trust region subproblems are of great interests because they are important parts of the trust region algorithms for nonlinear optimization. To construct efficient algorithms for solving these subproblems

In this paper we review some of the widely used trust region subproblems and some matrix computation problems related to these trust region subproblems.

2. TRS subproblem

In this section, we consider the solutions and approximate solutions of the trust region subproblem(TRS) which has the following form:

$$\min_{d \in \Re^n} \quad g^T d + \frac{1}{2} d^T B d = \phi(d) \tag{2.1}$$

$$\text{s. t.} \quad ||d||_2 \leq \Delta, \tag{2.2}$$

where $\Delta > 0$, $g \in \Re^n$, and $B \in \Re^{n \times n}$ is symmetric. Problem TRS (2.1)-(2.2) is a subproblem of trust region algorithms for unconstrained optimization.

The following lemma is well known (for example, see [14] and [15]):

Lemma 2.1 *A vector $d^* \in \Re^n$ is a solution of (2.1)-(2.2) if and only if there exists $\lambda^* \geq 0$ such that*

$$(B + \lambda^* I)d^* = -g \tag{2.3}$$

and that $B + \lambda^* I$ is positive semi-definite, $||d^*||_2 \leq \Delta$ and

$$\lambda^*(\Delta - ||d^*||_2) = 0. \tag{2.4}$$

It is easy to see from the above lemma that to solve the trust region subproblem TRS (2.1)-(2.2) is equivalent to find the correct parameter λ^* and solve the linear system (2.3). Therefore we can easily see that TRS is closed related to matrix computation problems. Indeed, we will see that an approximate solution of subproblem (2.1)-(2.2) can be computed by solving one of more systems of linear equations having the form (2.3).

Let d^* be a solution of problem (2.1)-(2.2) and λ^* be the multiplier satisfying conditions in the above lemma. If $B + \lambda^* I$ is positive definite, then d^* is uniquely defined by

$$d^* = -(B + \lambda^* I)^{-1} g. \tag{2.5}$$

The case where $B + \lambda^* I$ has zero eigenvalues is called "hard case". In this case, relation (2.3) implies that g is in the range space of $B + \lambda^* I$ and d^* can be written in the form:

$$d^* = -(B + \lambda^* I)^{+} g + v, \tag{2.6}$$

where v is a vector in the null space of $B + \lambda^* I$. On other hand, if g is in the range space of $B + \lambda^* I$ then any vector d^* given by (2.6) is also a solution of (2.1)-(2.2) provided that $||d^*||_2 \leq \Delta$ and that $\lambda^*(\Delta - ||d^*||_2) = 0$.

Unless in the hard case, λ^* is also the unique solution of the following equation

$$\psi(\lambda) = \frac{1}{||(B + \lambda I)^{-1} g||_2} - \frac{1}{\Delta} = 0. \tag{2.7}$$

Function $\psi(\lambda)$ is well defined for $\lambda \in (-\sigma_n(B), +\infty)$, where $\sigma_n(B)$ is the least eigenvalue of B. $\psi(\lambda)$ is concave and strictly monotonically increasing in $(-\sigma_n(B), +\infty)$ (For example, see [11]). In fact, the first order and second order derivatives of $\psi(\lambda)$ can be easily computed, thus Newton's method can be used to calculate λ^*. The Newton's iteration is

$$\begin{aligned} \lambda_+ &= \lambda - \frac{\psi(\lambda)}{\psi'(\lambda)} \\ &= \lambda - \frac{g^T (B + \lambda I)^{-3} g}{||(B + \lambda I) g||_2^3} \left[\frac{1}{||(B + \lambda I)^{-1} g||_2} - \frac{1}{\Delta} \right]. \end{aligned} \tag{2.8}$$

Based on Newton's iteration (2.8), numerical algorithms for problem (2.1)-(2.2) have been given by [14] and [15].

In the hard case, we have that

$$\lambda^* = -\sigma_n(B), \tag{2.9}$$

where $\sigma_n(B)$ is the least eigenvalue of B. If $-\sigma_n(B) = 0$, we can easily see that $-B^+g$ is a solution of problem (2.1)-(2.2). Hence the "real" hard case is that (2.9) is satisfied and $\sigma_n(B) < 0$. For any $\lambda \in (-\sigma_n(B), +\infty)$, Newton's step will normally make the matrix $B + \lambda_+ I$ have negative eigenvalue. Hence Newton's step (2.8) can only be used to adjust the lower bound λ_L. Based on these observations, we suggest to use the Newton's step for an equivalent equation

$$\tilde{\psi}(\mu) = \psi(\frac{1}{\mu}) = 0. \tag{2.10}$$

The numerical methods based on Newton's method for (2.7) needs to compute the Cholesky factorization of $(B + \lambda I)$, which is not desirable especially when B is a large sparse matrix.

Now, we discuss the conjugate gradient method for problem (2.1)-(2.2). The conjugate gradient method for minimize the convex function

$$\phi(d) = g^T d + \frac{1}{2} d^T B d \tag{2.11}$$

is iterative and it generates the iterates by the following formulae:

$$x_{k+1} = x_k + \alpha_k d_k \tag{2.12}$$

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{2.13}$$

where $g_k = \nabla\phi(x_k)$, and

$$\alpha_k = \frac{-d_k^T g_k}{d_k^T B d_k} \tag{2.14}$$

$$\beta_k = \frac{||g_{k+1}||^2}{-d_k^T g_k}, \tag{2.15}$$

with $x_1 = 0$ and $d_1 = -g$.

The conjugate gradient method has the nice finite termination property which mean that $x_k = -B^{-1}g$ for some $k \leq n+1$ if B is positive definite.

Steihaug[23] was the first to use the conjugate gradient method to solve the general trust region subproblem (2.1)-(2.2). Even without assuming the positive definite of B, we can continue the conjugate gradient method provided that $d_k^T B d_k$ is positive. If the iterate $x_k + \alpha_k d_k$ computed is in the trust region ball, it can be accepted, and the conjugate gradient iterates can be continued to the next iteration. Whenever $d_k^T B d_k$ is not positive or $x_k + \alpha_k d_k$ is outside the trust region, we can take the longest step along d_k within the trust region and terminate the calculations.

Algorithm 2.2 *(Truncated Conjugate Gradient Method For Trust Region Subproblem)*

Step 0 *Given* $g \in \Re^n$, $B \in \Re^{n\times n}$ *symmetric;*
$x_1 = 0,\ g_1 = g,\ d_1 = -g,\ k = 1.$

Step 1 *If* $||g_k|| = 0$ *then set* $x^* = x_k$ *and stop;*
Compute $d_k^T B d_k$*; if* $d_k^T B d_k^T \leq 0$ *then go to Step 3;*
Calculate α_k *by (2.14).*

Step 2 *If* $||x_k + \alpha_k d_k|| \geq \Delta$ *then go to Step 3;*
Set x_{k+1} *by (2.12) and* $g_{k+1} = g_k + \alpha_k B d_k$*;*
Compute β_k *by (2.15) and set* d_{k+1} *by (2.13);*
$k := k+1$*, go to Step 1.*

Step 3 *Compute* $\alpha_k^* \geq 0$ *satisfying* $||x_k + \alpha_k^* d_k|| = \Delta$*;*
Set $x^* = x_k + \alpha_k^* d_k$*, and Stop.*

The solution obtained by the above modified conjugate gradient method can satisfy the sufficient descent condition.

Lemma 2.3 *Let* x^* *computed by Algorithm 2.2, we have that*

$$\phi(0) - \phi(x^*) \geq \frac{1}{2}||g|| \min\{\Delta, ||g||/||B||\}. \tag{2.16}$$

Condition (2.16) plays an important role in the convergence analysis of trust region algorithms(see, for example, [18]).

We believe that if B is positive definite, the solution obtained by Algorithm 2.2 will yield a reduction in the object quadratic function at least half of the maximum reduction that can be obtained in the trust region. Namely, we believe that the following conjecture is true.

Conjecture 2.4 *Let* x^* *be computed by Algorithm (2.2), and let* d^* *be the solution of (2.1)-(2.2), if* B *is positive definite, then*

$$\phi(0) - \phi(x^*) \geq \frac{1}{2}[\phi(0) - \phi(d^*)]. \tag{2.17}$$

We have tested some randomly generated problems which show that our conjecture is likely to be true. However we have not yet been able to prove or disprove our conjecture theoretically.

If the corresponding Lagrange multiplier λ^* are known, the solution of (2.1)-(2.2) can be obtained by applying the conjugate gradient method directly to the linear system $(B + \lambda^*)d = -g$. However, in practice we do not know the value of λ^* before the problem (2.1)-(2.2) is solved. The following algorithm is a slightly modification of Algorithm 2.2 which tries to solve $(B + \lambda I)d = -g$ by the conjugate gradient method which modifies the parameter $\lambda \geq 0$ automatically. The main technique for updating the parameter λ is simple. When the conjugate gradient step is close to the boundary of the trust region, the parameter λ is increased.

Algorithm 2.5 *(Modified Conjugate Gradient Methods for TRS)*

Step 0 Given $g \in \Re^n$, $B \in \Re^{n\times n}$ symmetric;
$x_1 = 0$, $g_1 = g$, $d_1 = -g$, $\lambda = 0$, $\epsilon > 0$ very small, $k = 1$.

Step 1 If $\|x_k\| \geq \Delta - \epsilon$ then stop;
If $||g_k|| > 0$ go to Step 2;
If B positive definite then stop;
Find d_k such that $d_k^T B d_k < 0$ and $d_k^T x_k \geq 0$;
Go to Step 4;

Step 2 Compute $d_k^T B d_k$; if $d_k^T B d_k \leq 0$ then go to Step 4;
Calculate α_k by (2.14).

Step 3 If $||x_k + \alpha_k d_k|| \geq 0.5(\|x_k\| + \Delta)$ then go to Step 4;
Set x_{k+1} by (2.12) and $g_{k+1} = g_k + \alpha_k B d_k$;
Compute β_k by (2.15) and set d_{k+1} by (2.13);
$k := k + 1$, go to Step 1.

Step 4 Compute $\alpha_k^* \geq 0$ satisfying $||x_k + \alpha_k^* d_k|| = 0.5(\|x_k\| + \Delta)$;
Compute

$$\lambda = (-d_k^T g_k/\alpha^* - d_k^T B d_k)/\|d_k\|^2. \tag{2.18}$$

Set $B := B + \lambda I$;
Set $x_{k+1} = x_k + \alpha_k^* d_k$, and $g_{k+1} = g_k + \alpha_k^* B d_k$;
$d_{k+1} = -g_{k+1}$, $k := k + 1$; go to Step 1.

The above algorithm trys to find an approximate solution to the system $(B+\lambda I)d = -g$ by minimizing $g^T d + 0.5 d^T(B+\lambda I)d$ by the conjugate gradient method with the parameter λ updated automatically.

Assume that B is positive definite. If Newton's step $d = -B^{-1}g$ is in the trust region, it is easy to see that Newton's step is the solution of the problem (2.1)-(2.2). Therefore we can see that Newton's step is the solution of the trust region subproblem when the trust region bound is sufficienly large. On the other hand, if the trust region bound is very small, we can easily see that the solution will be very close to the steepest descent direction. Therefore, it is natural to consider to obtain an approximate solution of (2.1)-(2.2) by solving

$$\min_{d \in S} \quad g^T d + \frac{1}{2} d^T B d \tag{2.19}$$

$$\text{s. t.} \quad ||d|| \leq \Delta, \tag{2.20}$$

where

$$S = Span\{-g, -B^{-1}g\}. \tag{2.21}$$

It is easy to shown that the solution of (2.19)-(2.20) satisfies the sufficient descent condition (2.16) because $Spang \subset S$.

But, we have the following negative result about the 2-dimensional optimal step.

Lemma 2.6 *Let d^* be the solution of (2.1)-(2.2) and s^* be the solution of (2.19)-(2.20), assume that B is positive definite. Let $cond(B)$ be the condition number of B which is the ration between the largest and smallest eigenvalue of B, if*

$$\lim cond(B) = +\infty, \tag{2.22}$$

then

$$\lim \frac{\phi(0) - \phi(s^*)}{\phi(0) - \phi(d^*)} = 0. \tag{2.23}$$

Proof Consider the following example. Let $n = 3$, $g = (-1 \quad -\epsilon \quad -\epsilon^3)^T$, and

$$B = \begin{pmatrix} \epsilon^{-3} & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & \epsilon^3 \end{pmatrix}, \tag{2.24}$$

where $\epsilon > 0$ is a very small positive number.

It is easy to see that the Newton's step

$$-B^{-1}g = \begin{pmatrix} \epsilon^3 \\ \epsilon \\ 1 \end{pmatrix} \tag{2.25}$$

is the solution if $\Delta > 1+\epsilon+\epsilon^3$. If $\Delta < 1$, the minimizer of problem (2.1)-(2.2) can be written as

$$d(\lambda) = \begin{pmatrix} \frac{\epsilon^3}{1+\epsilon^3\lambda} \\ \frac{\epsilon}{\lambda+1} \\ \frac{\epsilon^3}{\lambda+\epsilon^3} \end{pmatrix}, \tag{2.26}$$

for some $\lambda > 0$ such that $||d(\lambda)|| = \Delta$. Specifically, if we let $\Delta = \epsilon/2$, the condition $||d(\lambda)|| = \Delta$ gives that

$$\lambda = 1 + O(\epsilon), \tag{2.27}$$

Therefore the maximum reduction in the trust region is

$$\phi(0) - \phi(d(\lambda)) = -\frac{1}{2}d(\lambda)^T g = \frac{\epsilon^2}{2} + O(\epsilon^3). \tag{2.28}$$

Now we consider the minimizer in the 2-dimensional subspace spanned by g and $B^{-1}g$. The solution can be written as

$$\begin{aligned} \bar{d}(\bar{\lambda}) &= -(g \quad B^{-1}g)\left[(g \quad B^{-1}g)^T(B+\bar{\lambda}I)(g \quad B^{-1}g)\right]^{-1}(g \quad B^{-1}g)^T g \\ &= -\begin{pmatrix} 1 & \epsilon^3 \\ \epsilon & \epsilon \\ \epsilon^3 & 1 \end{pmatrix}\left[\begin{pmatrix} \epsilon^{-3}+\epsilon^2+\epsilon^9 & 1+\epsilon^2+\epsilon^6 \\ 1+\epsilon^2+\epsilon^6 & \epsilon^2+2\epsilon^3 \end{pmatrix}\right. \\ &\qquad \left. +\bar{\lambda}\begin{pmatrix} 1+\epsilon^2+\epsilon^6 & \epsilon^2+2\epsilon^3 \\ \epsilon^2+2\epsilon^3 & 1+\epsilon^2+\epsilon^6 \end{pmatrix}\right]^{-1}\begin{pmatrix} 1+\epsilon^2+\epsilon^6 \\ \epsilon^2+2\epsilon^3 \end{pmatrix}. \end{aligned} \tag{2.29}$$

The requirement $||\bar{d}(\bar{\lambda})|| = \Delta$ implies that

$$\bar{\lambda} = 2\epsilon + O(\epsilon^2). \tag{2.30}$$

Therefore it follows the maximum reduction in the 2-dimensional subspace within the trust region is

$$\phi(0) - \phi(\bar{d}(\bar{\lambda})) = -\frac{1}{2}\bar{d}(\bar{\lambda})^T g = O(\epsilon^3). \tag{2.31}$$

Now relations (2.28) and (2.31) indicate the lemma is true. □

The above lemma shows that even though the 2-dimension minimizer satisfies the sufficient descent condition and such a subspace minimizer and similar approximate solutions such as dog-leg step or double dog-leg step are widely used in practice(for example, see [21],[10], [18] and [22]), it is possible that such inexact solutions yield very small reduction in the objective function comparing to the maximum deduction in the objective function in the whole trust region.

Recently, there are many research using semi-definite programming techniques to study subproblem (2.1)-(2.2). Such approaches normally require to find the least eigenvalues of the enlarged matrices having the form

$$\begin{bmatrix} t & g^T \\ g & B \end{bmatrix} , \tag{2.32}$$

where t is a parameter, more details can be found in Rendl and Wolkowicz[20] and Xin Chen [3]. Lanczos method can be used to compute the smallest eigenvalue of (2.32), which does not need to calculate matrix factorizations.

3. CDT subproblem

For equality constrained optimization, the linearized constraints are a system of linear equations. The system may have no feasible point within the trust region. One way to handle this difficulty is to replace the linear equations by a single constraint which imposes an upper bound to the sum of squares of the linearized constraints. This gives a trust region subproblem in the following form.

$$\min \phi(d) = g^T d + \frac{1}{2} d^T B d \tag{3.1}$$

$$s.t. \quad \|A^T d + c\|^2 \leq \xi^2 \tag{3.2}$$

$$\|d\|^2 \leq \Delta^2, \tag{3.3}$$

where $g \in \Re^n$, $B \in \Re^{n\times n}$ symmetric, $A \in \Re^{n\times m}$, $\xi \geq 0$ and $\Delta > 0$. Subproblem (3.1)-(3.2) was first proposed by Celis, Dennis and Tapia[2], hence it is usually called the CDT problem. This subproblem is also used in a trust region algorithm by Powell and Yuan[19] If the parameter

$$\xi = \xi_{min} = \min_{\|d\|\leq\Delta} \|A^T d + c\|, \tag{3.4}$$

it can be shown that the CDT problem is equivalent to a TRS subproblem in the null space of A^T, which can be solved by the methods discussed in the previous section. Therefore, it is without loss of generality to assume that

$$\xi \geq \xi_{min}. \tag{3.5}$$

The optimality condition for the CDT problem is given by Yuan[24].

Theorem 3.1 *Let d^* be a global solution of the problem (3.1)-(3.3). Assume that (3.5) holds, there exist Lagrange multipliers $\lambda^* \geq 0$ and $\mu^* \geq 0$ such that*

$$(B + \lambda^* I + \mu^* AA^T)d^* = -(g + \mu^* Ac) \tag{3.6}$$

and the complementarity conditions

$$\lambda^*[\Delta - ||d^*||_2] = 0 \tag{3.7}$$

$$\mu^*[\xi - ||c + A^T d^*||_2] = 0 \tag{3.8}$$

are satisfied. Furthermore, if the multipliers λ^*, μ^* are unique, then the matrix

$$H(\lambda^*, \mu^*) = B + \lambda^* I + \mu^* AA^T \tag{3.9}$$

has at most one negative eigenvalue.

From the above theorem, we can see that the CDT problem is closely related to find a solution of the linear system (3.6). An important issue is to find the correct Lagrange multipliers λ^* and μ^*. The above result is a necessary condition. For sufficient condition, we have the following result.

Theorem 3.2 *If d^* is a feasible point of (3.2)-(3.3), if λ^* and μ^* are two non-negative numbers such that (3.6)-(3.8) hold, and if the matrix $H(\lambda^*, \mu^*)$ is positive semi-definite, then d^* is a global solution of (3.1)-(3.3).*

We can easily see that there is a gap between the necessary conditions and the sufficient conditions. In the case of the TRS problem discussed in the previous section, the necessary conditions and sufficient conditions coincide. But, for the CDT problem, it is known that the matrix $H(\lambda^*, \mu^*)$ may have a negative eigenvalue when one of the constraints is inactive, and it may have more than one negative eigenvalue if d^* and $A(c + A^T d^*)$ are linearly dependent. The possibility of the indefiniteness of $H(\lambda^*, \mu^*)$ may lead to numerical difficulties when we trying to find a solution of the CDT problem by solving (3.6).

If B is positive semidefinite, the necessary conditions and sufficient conditions are the same. If B is positive definite, the CDT problem (3.1)-(3.3) can be solved by solving its dual problem. The dual problem for (3.1)-(3.3) is

$$\max_{(\lambda,\mu)\in\Re^2_+} \Psi(\lambda, \mu), \tag{3.10}$$

where

$$\Psi(\lambda, \mu) = \phi(d(\lambda, \mu)) + \frac{1}{2}\lambda(||d(\lambda, \mu)||_2^2 - \Delta^2) + \frac{1}{2}\mu(||c + A^T d(\lambda, \mu)||_2^2 - \xi^2) \tag{3.11}$$

and $d(\lambda, \mu)$ is defined by

$$d(\lambda, \mu) = -H(\lambda, \mu)^{-1}(g + \mu A c). \tag{3.12}$$

An algorithm based on Newton's method for the dual problem (3.10) is given by Yuan[25]. A step is truncated whenever it gives an infeasible (λ, μ). Line searches are also used to ensure convergence. It is also shown that the algorithm is quadratic convergent. More details can be found in Yuan (1991).

Zhang[27] gives an variable elimination algorithm for solving (3.14). For any $\mu \geq 0$, $\lambda(\mu)$ is defined to be the unique solution of

$$||d(\lambda, \mu)||_2 \leq \Delta, \quad \lambda(\Delta - ||d(\lambda, \mu)||_2) = 0, \ \lambda \geq 0. \tag{3.13}$$

It is easy to see that (3.13) is equivalent to the first row of the following system

$$\begin{pmatrix} \bar{\psi}(\lambda, \mu) \\ \hat{\psi}(\lambda, \mu) \end{pmatrix} \geq 0, \qquad (\lambda, \mu)^T \begin{pmatrix} \bar{\psi}(\lambda, \mu) \\ \hat{\psi}(\lambda, \mu) \end{pmatrix} = 0, \tag{3.14}$$

where

$$\bar{\psi}(\lambda, \mu) = \frac{1}{||d(\lambda, \mu)||_2} - \frac{1}{\Delta}, \tag{3.15}$$

and

$$\hat{\psi}(\lambda, \mu) = \frac{1}{||c + A^T d(\lambda, \mu)||_2} - \frac{1}{\xi}. \tag{3.16}$$

By the definition of $\lambda(\mu)$, (3.14) reduced to the following unvariable nonlinear system:

$$\hat{\psi}(\lambda(\mu), \mu) \geq 0, \quad \mu \hat{\psi}(\lambda(\mu), \mu) = 0, \ \mu \geq 0. \tag{3.17}$$

In the easy case that $\hat{\psi}(\lambda(0), 0) \geq 0$, it can be seen that $(\lambda(0), 0)$ is a solution. Otherwise, Zhang's algorithm applies Newton's method to solve

$$\hat{\psi}(\lambda(\mu), \mu) = 0. \tag{3.18}$$

When B is a general symmetric matrix, Sufficient conditions and necessary conditions are not the same, detailed discussions can be found in Chen and Yuan[6]. Recently Xiong-Da Chen[5] studied the structure of the dual space of the CDT problem. Some interesting results have been obtained on the location regions of the Lagrange multipliers corresponding to the global solution, which can be found in Chen and Yuan[7]. Chen[5] also considered the parameterized problem

$$\min \quad \phi(d) \tag{3.19}$$

$$s.t. \quad \omega(\|d\|^2 - \Delta^2) + (1 - \omega)(\|A^T d + c\|^2 - \xi^2) \leq 0, \tag{3.20}$$

where $\omega \in [0,1]$. The above subproblem is a single ball problem. Relations between the multiplier for (3.19)-(3.20) and the multipliers for (3.1)-(3.3) are discussed in Chen and Yuan[8].

A direct way for solving (3.1)-(3.3) is by applying a truncated conjugate gradient method or by a projected conjugate gradient method similar to we discussed in the previous section.

Algorithm 3.3 *(Truncated Conjugate Gradient Methods for CDT)*

Step 0 *Given* $g \in \Re^n$, $B \in \Re^{n\times n}$ *symmetric;*
$c \in \Re^m$, $A \in \Re^{n\times m}$, $\Delta > 0$, $\xi > 0$ *;*
Given an interior point x_1 *of (3.2)-(3.3);*
$d_1 = g_1 = \phi(x_1)$*;* $k = 1$.

Step 1 *If* $||g_k|| = 0$ *then set* $x^* = x_k$ *and stop;*
Compute $d_k^T B d_k$*; if* $d_k^T B d_k^T \le 0$ *then go to Step 3;*
Calculate α_k *by (2.14).*

Step 2 *If* $x_k + \alpha_k d_k$ *is in the interior of (3.2)-(3.3) then go to Step 3;*
Set x_{k+1} *by (2.12) and* $g_{k+1} = g_k + \alpha_k B d_k$*;*
Compute β_k *by (2.15) and set* d_{k+1} *by (2.13);*
$k := k+1$*, go to Step 1.*

Step 3 *Compute* $\alpha_k^* \ge 0$ *such that* $x_k + \alpha_k^* d_k$ *is on the boundary of (3.2)-(3.3);*
Set $x^* = x_k + \alpha_k^* d_k$*, and Stop.*

The algorithm is basically a conjugate gradient method starting from an interior point of the feasible region. It stops whenever the iterate reaches the boundary or outside the feasible region in which case the step is truncated and a boundary point is taken as the approximate solution. The algorithm can be slightly modified so that once the iterate point reaches the boundary, it searches along the boundary by using projected search directions.

A generalization of the CDT problem is to minimizer a quadratic function subject to two general quadratic constraints:

$$\min \quad q_1(x) \tag{3.21}$$
$$s.t. \quad q_2(x) \le 0 \tag{3.22}$$
$$q_3(x) \le 0, \tag{3.23}$$

where $q_i(x)(i = 1,2,3)$ are general functions in $\Re^n$. This problem was studied by Peng and Yuan [17]. A special case of (3.21)-(3.23) is the case when $q_i(x) =$

$x^T C_i x$, which gives the following problem

$$\begin{aligned} min \quad & x^T C_1 x & (3.24)\\ s.t. \quad & x^T C_2 x \le 0, & (3.25)\\ & x^T C_3 x \le 0, & (3.26) \end{aligned}$$

where $C_i (i = 1, 2, 3)$ are symmetric matrices in $\Re^{n\times n}$. If 0 solves (3.24)-(3.26), there exists $(\alpha_0, \beta_0) \in \Re^2$, (α_0, β_0) maximizes the least eigenvalue of $C_1+\alpha C_2+\beta C_3$, such that $C_1 + \alpha_0 C_2 + \beta_0 C_3$ has at most two negative eigenvalues. And when $\alpha C_2 + \beta C_3$ is indefinite for all $(\alpha, \beta) \in \Re^2((\alpha, \beta) \ne 0)$ and the least eigenvalue of $C_1 + \alpha_0 C_2 + \beta_0 C_3$ is negative, α_0, β_0 must be greater than 0. For more details, please see Peng and Yuan[17].

If 0 solves (3.24)-(3.26), it can be seen that we have

$$\max\{x^T C_1 x,\ x^T C_2 x,\ x^T C_3 x\} \ge 0, \tag{3.27}$$

for all $x \in \Re^n$. Yuan [24] gives a very interesting result about two quadratic forms. It reads as follows:

Theorem 3.4 *Let $C_1, C_2 \in \Re^{n\times n}$ be two symmetric matrices and A and B be two closed sets in $\Re^n$ such that*

$$A \cup B = \Re^n. \tag{3.28}$$

If we have

$$x^T C_1 x \ge 0, x \in A, x^T C_2 x \ge 0, x \in B, \tag{3.29}$$

then there exists a $t \in [0, 1]$ such that the matrix

$$tC_1 + (1 - t)C_2 \tag{3.30}$$

is positive semi-definite.

J.P.Crouzeix *et.al.* [9] pointed out that Yuan's result is actually an alternate theorem. They extended Theorem 3.4 to a locally convex topological linear space and showed that it can not be extended to more than two matrices and copositive matrices in a simple way.

Recently Chen and Yuan[4] showed that if (3.27) holds, there exist a convex linear combination of $C_i(i = 1, 2, 3)$ that has at most one negative eigenvalue. This result and Theorem 3.4 indicates that the following conjecture might be true.

Conjecture 3.5 *Let $C_i(i = 1, ..., m)$ be m symmetrical matrices in $\Re^{n\times n}$. If*

$$max_{1\le i\le m}\{x^T C_i x\} \ge 0, \text{ for every } x \in \Re^n. \tag{3.31}$$

Then there exists a $C = \sum_{i=1}^m t_i C_i, (\sum_{i=1}^m t_= 1, t_i \ge 0, i = 1, ..., m)$, such that C has at most $m - 2$ negative eigenvalue.

References

[1] R. Byrd, R.B. Schnabel and G.A. Shultz, "A trust region algorithm for nonlinearly constrained optimization", *SIAM J. Numer. Anal.* 24(1987) 1152-1170.

[2] 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. Schnabel, eds., *Numerical Optimization* (SIAM, Philadelphia, 1985) pp. 71-82.

[3] Xin Chen, *Trust Region Subproblems and Extensions*, Master Thesis, Institute of Computational Mathematics and Scientific/ Engineering Computing, Chinese Academy of Sciences, 1998.

[4] Xin Chen and Y. Yuan, "A note on quadratics forms", Research Report, Institute of Computational Mathematics and Scientific/ Engineering Computing, Chinese Academy of Sciences, 1997.

[5] Xiong Da Chen, *On the Subproblem of Trust Region Algorithms*, Ph. D. Thesis, Institute of Computational Mathematics and Scientific/ Engineering Computing, Chinese Academy of Sciences, 1998.

[6] X. D. Chen and Y. Yuan, "Optimality conditions for CDT subproblem", Research Report, Institute of Computational Mathematics and Scientific/ Engineering Computing, Chinese Academy of Sciences, 1997.

[7] X. D. Chen and Y. Yuan, "On local solutions of CDT subproblem", Institute of Computational Mathematics and Scientific/ Engineering Computing, Chinese Academy of Sciences, 1998.

[8] X. D. Chen and Y. Yuan, "On maxima of dual function of CDT subproblem", Institute of Computational Mathematics and Scientific/ Engineering Computing, Chinese Academy of Sciences, 1998.

[9] J. P. Crouzeix, J. E. Martínez-Legaz and A. Seeger, *An alternative theorem for quadratic forms and extensions,* Preprint No 98, Department of Math, University of Barcelona, 1991.

[10] J.E. Dennis and H.H.W. Mei, "Two new unconstrained optimization algorithms which use function and gradient values", *J. Opt. Theory and Applns.* 28(1979) 453-482.

[11] J.E. Dennis Jr. and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1983)

[12] R. Fletcher, "A model algorithm for composite NDO problem", *Math. Prog. Study* 17(1982) 67-76.

[13] R. Fletcher, "Second order correction for nondifferentiable optimization", in: G.A. Watson, ed., *Numerical Analysis* (Springer-Verlag, Berlin, 1982) pp. 85-115.

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

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

[16] E. O. Omojokun, *Trust Region Algorithms for Optimization with Nonlinear Equality and Inequality Constraints*, Ph. D. Thesis, University of Colorado at Boulder, 1989.

[17] J. Peng and Y. Yuan, "Optimality Conditions for the Minimization of a Quadratic with Two Quadratic Constraints", *SIAM J. Optimization* 7(1997) 574-594.

[18] M.J.D. Powell, "A new algorithm for unconstrained optimization", in: J.B. Rosen, O.L. Mangasarian and K. Ritter, eds., *Nonlinear Programming* (Academic Press, New York, 1970) pp. 31-66.

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

[20] "A semidefinite framework for trust region subproblems with applications to large scale minimization", *Math. Prog.* 77(1997) 273–299.

[21] G.A. Shultz, R.B. Schnabel and R.H. Byrd, "A family of trust-region-based algorithms for unconstrained minimization with strong global convergence", *SIAM J. Numer. Anal.* 22(1985) 47-67.

[22] D.C. Sorensen, "Newton's method with a model trust region modification", *SIAM J. Numer. Anal.* 20(1982) 409-426.

[23] T. Steihaug, "The conjugate gradient method and trust regions in large scale optimization", *SIAM J. Numer. Anal.* 20(1983) 626-637.

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

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

[26] Y. Yuan, "On the convergence of a new trust region algorithm", *Numerische Mathematik*, 70(1995) 515–539.

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