Math. Program., Ser. A 87: 561–573 (2000)

Digital Object Identifier (DOI) 10.1007/s101079900126

Y. Yuan

On the truncated conjugate gradient method*

Received January 19, 1999 / Revised version received October 1, 1999
Published online November 30, 1999 – © Springer-Verlag 1999

Abstract. In this paper, we consider the truncated conjugate gradient method for minimizing a convex quadratic function subject to a ball trust region constraint. It is shown that the reduction in the objective function by the solution obtained by the truncated CG method is at least half of the reduction by the global minimizer in the trust region.

Key words. unconstrained optimization – trust region – conjugate gradient

1. Introduction

Consider the unconstrained optimization problem

$$\min_{x\in\Re^n} f(x), \tag{1}$$

where $f(x)$ is continuously differentiable. Trust region algorithms for (1) often need to solve the following subproblem: (TRS)

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

subject to

$$\|d\| \le \Delta, \tag{3}$$

where $\Delta > 0$ is a trust region bound, $g \in \Re^n$ is the gradient of the objective function $f(x)$ at the current iterate, and $B \in \Re^{n\times n}$ symmetric is an approximation to the Hessian of $f(x)$. At each iteration of a trust region algorithm, a problem in the form of (2)–(3) has to be solved exactly or inexactly to obtain a trial step. The trial step, often called as the trust region step, will either be accepted or rejected after testing some test condition based on the predicted reduction and the actual reduction of the objective function. For more details, please see Fletcher [2].

The following lemma is well known (for example, see Gay [3] and More and Sorensen [4]):

Y. Yuan: State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100080, China. e-mail: yyx@lsec.cc.ac.cn

Mathematics Subject Classification (1991): 65K, 90C

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

Lemma 1. *A vector $d^* \in \Re^n$ is a solution of (2)–(3) if and only if there exists $\lambda^* \geq 0$ such that*

$$(B + \lambda^* I)d^* = -g \tag{4}$$

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

$$\lambda^*(\Delta - ||d^*||_2) = 0. \tag{5}$$

For a give trial step s, the prediction of the objective function is given by

$$Pred(s) = \phi(0) - \phi(s). \tag{6}$$

It is shown by Powell [5] that trust region algorithms for (1) is convergent if the trust region step satisfies

$$Pred(s) \geq c\|g\| \min\{\Delta, \|g\|/\|B\|\} \tag{7}$$

and some other conditions on B are satisfied. It is easy to see that

$$\phi(0) - \min_{d \in Span\{g\}, \|d\| \leq \Delta} \phi(d) \geq \frac{1}{2}\|g\| \min\{\Delta, \|g\|/\|B\|\}. \tag{8}$$

Therefore it is quite common that in practice the trial step at each iteration of a trust region method is computed by solving the trust region subproblem (2)–(3) inexactly. One way to compute an inexact solution of (2)–(3) was the truncated conjugate gradient method proposed by Toint [7] and Steihaug [6]. The aim of this paper is to show that if B is positive definite, the function reduction obtained by the truncated conjugate gradient method is at least half of the reduction obtained by the exact solution.

2. The truncated CG method

The conjugate gradient method for

$$\min_{d \in \Re^n} \phi(d) = g^T d + \frac{1}{2} d^T B d \tag{9}$$

generates a sequence as follows:

$$x_{k+1} = x_k + \alpha_k d_k, \tag{10}$$

$$d_{k+1} = -g_k + \beta_k d_k, \tag{11}$$

where $g_k = \nabla\phi(x_k) = g + Bx_k$ and

$$\alpha_k = -g_k^T d_k / d_k^T B d_k, \qquad \beta_k = \|g_{k+1}\|^2 / \|g_k\|^2, \tag{12}$$

with the initial values

$$x_1 = 0, \quad d_1 = -g_1 = -g. \tag{13}$$

It can be shown that the conjugate gradient method terminates after at most n iterations (see Fletcher [2]). That is, there exists a integer $\bar{k} \leq n + 1$ such that $g_{\bar{k}} = 0$.

Lemma 2. *For any $k \geq 1$ such that $g_k \neq 0$ we have that*

$$d_k = -\|g_k\|^2 \sum_{i=1}^{k} \frac{g_i}{\|g_i\|^2} \tag{14}$$

$$x_{k+1} = -\sum_{i=1}^{k} \frac{g_i}{\|g_i\|^2} \sum_{j=i}^{k} \alpha_j \|g_j\|^2. \tag{15}$$

Proof. By definition, $d_1 = -g_1$, which shows that (14) holds for $k = 1$. Assume it holds for $k = 1, ..., \bar{k}$. If $g_{\bar{k}+1} \neq 0$, it follows from (11) and (12) that

$$d_{\bar{k}+1} = -g_{\bar{k}+1} + \frac{\|g_{\bar{k}+1}\|^2}{\|g_{\bar{k}}\|^2} d_{\bar{k}} = -g_{\bar{k}+1} + \frac{\|g_{\bar{k}+1}\|^2}{\|g_{\bar{k}}\|^2} \left(-\|g_{\bar{k}}\|^2 \sum_{i=1}^{\bar{k}} \frac{g_i}{\|g_i\|^2} \right) \tag{16}$$

$$= -\|g_{\bar{k}+1}\|^2 \left(\frac{g_{\bar{k}+1}}{\|g_{\bar{k}+1}\|^2} + \sum_{i=1}^{\bar{k}} \frac{g_i}{\|g_i\|^2} \right) = -\|g_{\bar{k}+1}\|^2 \sum_{i=1}^{\bar{k}+1} \frac{g_i}{\|g_i\|^2}.$$

Thus, by induction, (14) is true for all $k \geq 1$ provided that $g_k \neq 0$.

From (13), (10) and (14), we have that

$$x_{k+1} = \sum_{j=1}^{k} \alpha_j d_j = -\sum_{j=1}^{k} \alpha_j \|g_j\|^2 \sum_{i=1}^{j} \frac{g_i}{\|g_i\|^2}$$

$$= -\sum_{i=1}^{k} \frac{g_i}{\|g_i\|^2} \sum_{j=i}^{k} \alpha_j \|g_j\|^2, \tag{17}$$

which shows that (15) holds.

□

Toint [7] and Steihaug [6] were the first to use the conjugate gradient method to solve the general trust region subproblem (2)–(3). 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 1. *(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 (12).

Step 2. *If* $||x_k + \alpha_k d_k|| \geq \Delta$ *then go to Step 3;*
Set x_{k+1} *by (10) and* $g_{k+1} = g_k + \alpha_k B d_k$*;*
Compute β_k *by (12) and set* d_{k+1} *by (11);*
$k := k + 1$*, 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.*

Let x^* be the inexact solution of (2)–(3) obtained by the above truncated CG method and d^* be the exact solution of (2)–(3). If $n = 2$, Yuan [10] shows that

$$\frac{\phi(0) - \phi(x^*)}{\phi(0) - \phi(d^*)} \geq \frac{1}{2}. \tag{18}$$

It was also conjectured by Yuan [10] that (18) is true for all n. Numerical tests given by Chen [1] support this conjecture. Recently Tseng [8] shows that

$$\frac{\phi(0) - \phi(x^*)}{\phi(0) - \phi(d^*)} \geq \frac{1}{3}. \tag{19}$$

The main result of this paper is establishing (18) for all $n \geq 1$. Inequality (18) presents a reason why the Steihaug-Toint CG method works so well in practice.

Let $\bar{q} = \max_{\|d\| \leq \Delta} \phi(d)$, we have

$$\bar{q} - \phi(0) \geq \phi(0) - \phi(d^*), \tag{20}$$

if B is positive semi-definite. In this case, (18) and the above inequality imply that

$$\frac{\bar{q} - \phi(x^*)}{\bar{q} - \phi(d^*)} \geq \frac{3}{4}. \tag{21}$$

This kind of inequality is also of interests in complexity analysis(for example, see Ye [9]).

3. Conjugate gradient path

For any given orthogonal matrix Q, we define $\bar{g} = Q^T g$, and $\bar{B} = Q^T B Q$, we can easily see that the conjugate gradient method applied to

$$\bar{\phi}(\bar{d}) = \bar{g}^T \bar{d} + \frac{1}{2} d^T \bar{B} \bar{d} \tag{22}$$

will generate the iterates $\bar{x}_k = Q^T x_k$, $\bar{g}_k = Q^T g_k$ and $\bar{d}_k = Q^T d_k$. Thus, the conjugate gradient method is invariant by orthogonal transformation. Since for any give $g \in \Re^n$ and symmetric matrix B, there exists a orthogonal matrix Q such that $Q^T g$ is parallel to the first coordinate direction and $Q^T B Q$ is a tridiagonal matrix. Therefore, without loss of generality, throughout the rest of this paper, we assume that

$$g = \|g\| \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{23}$$

$$B = \begin{pmatrix} u_1 & v_1 & 0 & \dots & 0 & 0 \\ v_1 & u_2 & v_2 & \dots & 0 & 0 \\ 0 & v_2 & u_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & u_{n-1} & v_{n-1} \\ 0 & 0 & 0 & \dots & v_{n-1} & u_n \end{pmatrix} \tag{24}$$

In the following we define the path given by the conjugate gradient method. For all $k \geq 1$ such that $g_k \neq 0$, we denote

$$x(t) = x_k + (t-k)(x_{k+1} - x_k), \qquad \forall t \in [k, k+1]. \tag{25}$$

Define

$$B_k = \begin{pmatrix} u_1 & v_1 & 0 & \dots & 0 & 0 \\ v_1 & u_2 & v_2 & \dots & 0 & 0 \\ 0 & v_2 & u_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & u_{k-1} & v_{k-1} \\ 0 & 0 & 0 & \dots & v_{k-1} & u_k \end{pmatrix}, \tag{26}$$

which is the submatrix of the first k rows and first k columns of B. We assume that B is positive definite, which implies that all B_k are also positive definite. It is easy to prove that

Lemma 3. *If* $\prod_{i=1}^{k} v_i \neq 0$, *then*

$$x_{k+1} = -\|g\| \begin{pmatrix} B_k^{-1} e_1 \\ 0 \end{pmatrix}. \tag{27}$$

And

$$g_{k+1} = (-1)^k e_{k+1} \|g\| \frac{\prod_{i=1}^{k} v_i}{Det(B_k)}. \tag{28}$$

Proof. x_{k+1} is the solution of

$$\min_{d \in S_k} g^T d + \frac{1}{2} d^T B d \tag{29}$$

where

$$S_k = Span\{g, Bg, B^2 g, \dots, B^{k-1} g\}. \tag{30}$$

From the facts that $g = \|g\| e_1$, B is tridiagonal and $v_i \neq 0 (i = 1, ..., k-1)$, we can easily see that

$$S_k = Span\{e_1, e_2, \dots, e_k\}. \tag{31}$$

It follows from (29)–(31) that for $i = 1, \ldots, k$

$$e_i^T g_{k+1} = 0, \tag{32}$$

which gives that

$$e_i^T (g + Bx_{k+1}) = 0 \tag{33}$$

for all $i = 1, ..., k$. The above equation and the fact that $x_{k+1} \in S_k$ show the validity of (27).

From (27), we have that

$$g_{k+1} = g + Bx_{k+1} = -\|g\| v_k e_1^T B_k^{-1} e_k e_{k+1}. \tag{34}$$

For $k = 1$, we have that

$$g_2 = -\|g\| v_1 e_1^T B_1^{-1} e_1 e_2 = -e_2 \|g\| \frac{v_1}{Det(B_1)}. \tag{35}$$

If $k > 1$, because B_k is tridiagonal, we can see that

$$e_1^T B_k^{-1} e_k = (-1)^{k-1} \frac{\prod_{i=1}^{k-1} v_i}{Det(B_k)}, \tag{36}$$

which, together with (34), gives

$$g_{k+1} = (-1)^k \|g\| e_{k+1} \frac{\prod_{i=1}^{k} v_i}{Det(B_k)}. \tag{37}$$

It follows from (37) and (35) that relation (28) holds for all $k \geq 1$.

□

From the above lemma, we can see that

$$\|g_{k+1}\|^2 = \|g\|^2 \frac{\prod_{i=1}^{k} v_i^2}{(Det(B_k))^2}. \tag{38}$$

Lemma 4. *$k < n$ is the integer such that*

$$g_{k+1} = 0 \tag{39}$$

if and only if that k is the smallest integer that

$$v_k = 0. \tag{40}$$

Proof. The lemma follows from (37) directly.

□

If $g_{k+1} = 0$ for $k < n$, we can consider the problem in the subspace $Span\{e_1, e_2, ..., e_k\}$. Hence, there is no loss of generality in assuming that $v_i \neq 0$ for all $i = 1, ..., n-1$.

Lemma 5. *Let $x(t)(t \geq 1)$ is the conjugate gradient path. If $v_i \neq 0$ $(i = 1, ..., k-1)$, we have that*

$$x(t) = -\sum_{i=1}^{k} \gamma_i(t) Sign\left(e_i^T g_i\right) e_i, \qquad \text{for all } t \in [1, k+1] \tag{41}$$

where $\gamma_i(t) = 0$ for $t \in [1, i]$ and

$$\gamma_i(t) = \frac{1}{\|g_i\|} \left(\sum_{j=i}^{[t]-1} \alpha_j \|g_j\|^2 + (t - [t])\alpha_{[t]} \|g_{[t]}\|^2 \right) \tag{42}$$

for $t \in [i, k+1]$, where $[t]$ is the largest integer that is not greater than t. Here $Sign(e_i^T g_i) = 1$ if $e_i^T g_i > 0$, otherwise $Sign(e_i^T g_i) = -1$.

Proof. By definition

$$Sign\left(e_i^T g_i\right) e_i = \frac{g_i}{\|g_i\|}. \tag{43}$$

It follows from (25), (14), (15), (10) and the fact that $\gamma_i(t) = 0$ for $i \geq t$ that

$$\begin{aligned}
x(t) &= x_{[t]} + (t - [t])\alpha_{[t]} d_{[t]} \\
&= -\sum_{i=1}^{[t]-1} \frac{g_i}{\|g_i\|^2} \sum_{j=i}^{[t]-1} \alpha_j \|g_j\|^2 + (t - [t])\alpha_{[t]} \left(-\|g_{[t]}\|^2 \sum_{i=1}^{[t]} \frac{g_i}{\|g_i\|^2} \right) \\
&= -\sum_{i=1}^{[t]} \frac{g_i}{\|g_i\|^2} \left(\sum_{j=i}^{[t]-1} \alpha_j \|g_j\|^2 + (t - [t])\alpha_{[t]} \|g_{[t]}\|^2 \right) \\
&= -\sum_{i=1}^{[t]} \frac{g_i}{\|g_i\|} \gamma_i(t) \\
&= -\sum_{i=1}^{k} \frac{g_i}{\|g_i\|} \gamma_i(t).
\end{aligned} \tag{44}$$

This completes our proof.

□

The following corollaries are useful in our analysis.

Corollary 1. *If $v_i \neq 0$ $(i = 1, ..., k-1)$, then for each given integer $i \in [1, k]$ $\gamma_i(t)$ is strictly increasing for $t \in [i, k+1]$. Furthermore, we have that*

$$-x(t)^T g = \gamma_1(t)\|g\| \tag{45}$$

$$\|x(t)\|^2 = \sum_{i}^{k} (\gamma_i(t))^2, \tag{46}$$

for all $t \in [1, k+1]$. Thus, $-x(t)^T g$ and $\|x(t)\|$ are strictly monotonically increasing functions of t in $[1, k+1]$.

Proof. The increasing of $\gamma_i(t)$ can be directly shown from (42). (45) and (46) are consequences of (41).

□

Corollary 2. *If $v_i \neq 0$ $(i = 1, ..., k-1)$, then for each given integer $i \in [1, k]$ the relation*

$$\|g_i\|[\gamma_i(\hat{t}) - \gamma_i(\bar{t})] = \|g\|[\gamma_1(\hat{t}) - \gamma_1(\bar{t})] \tag{47}$$

holds for any two $\bar{t}$ and $\hat{t} \in [i, k+1]$.

Proof. (47) follows from (42).

□

Corollary 3. *If $v_i \neq 0$ $(i = 1, ..., k-1)$, then*

$$x(t)^T g(x(t)) \leq 0, \tag{48}$$

for all $t \in [1, k+1]$.

Proof. For any $t \in [1, k+1]$, there exists an integer $i \in [1, k]$ and a real number $\delta \in [0, 1]$ such that

$$t = \alpha i + (1-\alpha)(i+1). \tag{49}$$

Therefore we have that

$$g(x(t)) = \alpha g_i + (1-\alpha) g_{i+1}. \tag{50}$$

The above relation and (41) gives that

$$x(t)^T g(x(t)) = -\alpha\gamma_i(t)\|g_i\| - (1-\alpha)\gamma_{i+1}(t)\|g_{i+1}\| \leq 0. \tag{51}$$

This indicates that the corollary holds.

□

.

4. Conjugate gradient path of the exact solution of TRS

It follows from Lemma 1 that the exact solution d^* of the trust region subproblem (2)–(3) is the minimizer of $\phi(d) + \lambda^*\|d\|^2/2$. Thus d^* is the end of the conjugate gradient path if the objective function is $\phi(d) + \lambda^*\|d\|^2/2$.

For any given $\lambda > 0$, we consider the conjugate gradient method applied to the problem

$$\min_{d \in \Re^n} \phi(d, \lambda) = g^T d + \frac{1}{2} d^T (B + \lambda I) d. \tag{52}$$

Let the iterate points and the gradients generated be denoted by $x_i(\lambda)(i = 1, ..., n+1)$ and $g_i(\lambda)(i = 1, ..., n+1)$ respectively. We have that $x_1(\lambda) = x_1 = 0$ and $g_1(\lambda) = g_1 = g$. It follows from (28) that

$$g_{k+1}(\lambda) = \frac{Det(B_k)}{Det(B_k + \lambda I)}\, g_{k+1} \tag{53}$$

for $k \geq 1$. Thus, the conjugate gradient path is now given by

$$x(t, \lambda) = -\sum_{i=1}^{k} \gamma_i(t, \lambda) Sign\big(e_i^T g_i\big) e_i, \qquad \text{for all } t \in [1, k+1], \tag{54}$$

where $\gamma_i(t, \lambda) = 0$ for $t \in [1, i]$ and

$$\gamma_i(t, \lambda) = \frac{1}{\|g_i(\lambda)\|} \left(\sum_{j=i}^{[t]-1} \alpha_j(\lambda) \|g_j(\lambda)\|^2 + (t - [t]) \alpha_{[t]}(\lambda) \|g_{[t]}(\lambda)\|^2 \right) \tag{55}$$

for $t \in [i, k+1]$.

□

Lemma 6. *For any $\lambda > 0$ and $k \geq 1$ such that $g_k \neq 0$, we have*

$$-x_{k+1}(\lambda)^T g < -x_{k+1}^T g \tag{56}$$

$$\|x_{k+1}(\lambda)\| < \|x_{k+1}\|. \tag{57}$$

Furthermore,

$$\|g_k(\lambda)\| < \|g_k\| \tag{58}$$

if $k > 1$.

Proof. It follows from (27) that

$$-x_{k+1}(\lambda)^T g = \|g\|^2 e_1^T (B_k + \lambda I)^{-1} e_1 < \|g\|^2 e_1^T B_k^{-1} e_1 = -x_{k+1}^T g, \tag{59}$$

which shows that (56) is true.

Again, from (27), we have that

$$\|x_{k+1}(\lambda)\|^2 = \|g\|^2 e_1^T (B_k + \lambda I)^{-2} e_1 < \|g\|^2 e_1^T B_k^{-2} e_1 = \|x_{k+1}\|^2, \tag{60}$$

which gives (57). If $k > 1$, From (38) we have that

$$\|g_k(\lambda)\| = \|g_k\| \frac{Det(B_{k-1})}{Det(B_{k-1} + \lambda I)} < \|g_k\|. \tag{61}$$

This completes the proof of the lemma.

□

Theorem 1. *For any $\lambda > 0$ and $k \geq 1$ such that $g_k \neq 0$, there exists $t_k \in [1, k+1)$ such that*

$$-x_{k+1}(\lambda)^T g = -x(t_k)^T g, \tag{62}$$

and

$$\gamma_i(t_k) < \gamma_i(k+1, \lambda), \tag{63}$$

for all $i = 2, ..., k$.

Proof. If $k = 1$, we have

$$-x_{k+1}(\lambda)^T g = -x_2(\lambda)^T g = \alpha_1(\lambda) g^T g = \frac{\|g\|^4}{g^T(B+\lambda I)g} \tag{64}$$

and

$$-x(t)^T g = -(t-1)x_2^T g = (t-1)\frac{\|g\|^4}{g^T B g} \tag{65}$$

for $t \in [1, 2]$. Thus, (62) is true for $k = 1$ if we let $t_1 = 1 + g^T Bg/g^T(B+\lambda I)g$.

For the case when $k = 2$. If $g_2 \neq 0$, if follows from Corollary 1 that there exists $t_2 \in [1, 3)$ such that (62) holds. If $t_2 \leq 2$ then $\gamma_2(t) = 0$ which implies (63). Otherwise, we have that $t_2 \in (2, 3)$. It follows (47) and (58) that

$$\begin{aligned}
\gamma_2(t_2) = \gamma_2(t_2) - \gamma_2(2) &= \frac{\|g\|}{\|g_2\|}[\gamma_1(t_2) - \gamma_1(2)] \\
&< \frac{\|g\|}{\|g_2\|}[\gamma_1(t_2) - \gamma_1(t_1)] = \frac{\|g\|}{\|g_2\|}[\gamma_1(3,\lambda) - \gamma_1(2,\lambda)] \\
&= \frac{\|g\|}{\|g_2\|}\frac{\|g_2(\lambda)\|}{\|g(\lambda)\|}[\gamma_2(3,\lambda) - \gamma_2(2,\lambda)] \\
&< \gamma_2(3,\lambda) - \gamma_2(2,\lambda) = \gamma_2(3,\lambda),
\end{aligned} \tag{66}$$

which shows that (63) holds for $k = 2$.

Now we prove the theorem by induction. we assume that for some $k \geq 2$ there exists $t_k \in [1, k+1)$ such that (62) and (63) hold. If $g_{k+1} \neq 0$, from the above lemma and the monotone property of $x(t)^T g$, there exists a unique $t_{k+1} \in (t, k+2)$ such that

$$-x_{k+2}(\lambda)^T g = -x(t_{k+1})^T g. \tag{67}$$

Relation (47) implies that

$$\frac{\gamma_i(k+2,\lambda) - \gamma_i(k+1,\lambda)}{\gamma_1(k+2,\lambda) - \gamma_1(k+1,\lambda)} = \frac{\|g\|}{\|g_i(\lambda)\|}. \tag{68}$$

for all $i = 1, ..., k+1$.

On the other hand, let $L = [t_k]$ and $M = [t_{k+1}]$, we have that

$$\gamma_i(t_{k+1}) = \gamma_i(t_k) + \frac{1}{\|g_i\|}\left((L+1-t_k)\alpha_L\|g_L\|^2 + \sum_{j=L+1}^{M-1}\alpha_j\|g_j\|^2 + (t_{k+1}-M)\alpha_M\|g_M\|^2\right) \tag{69}$$

for all $i = 1, 2, ..., L$, and

$$\gamma_i(t_{k+1}) = \gamma_i(t_k) + \frac{1}{\|g_i\|}\left(\sum_{j=i}^{M-1}\alpha_j\|g_j\|^2 + (t_{k+1}-M)\alpha_M\|g_M\|^2\right) \tag{70}$$

for $i = L+1, ..., M$. Therefore if follows from the above two relations that

$$\frac{\gamma_i(t_{k+1}) - \gamma_i(t_k)}{\gamma_1(t_{k+1}) - \gamma_1(t_k)} = \frac{\|g_1\|}{\|g_i\|} \tag{71}$$

for $i = 1, ..., L$, and

$$\frac{\gamma_i(t_{k+1}) - \gamma_i(t_k)}{\gamma_1(t_{k+1}) - \gamma_1(t_k)} < \frac{\|g_1\|}{\|g_i\|} \tag{72}$$

for $i = L+1, ..., M$. Therefore, for $i = 2, ..., \min\{M, k\}$ it follows from (62), (63), (71) and (72) that

$$\begin{aligned}
\gamma_i(t_{k+1}) &= \gamma_i(t_k) + [\gamma_i(t_{k+1}) - \gamma_i(t_k)] \\
&\le \gamma_i(t_k) + \frac{\|g\|}{\|g_i\|}[\gamma_1(t_{k+1}) - \gamma_1(t_k)] \\
&= \gamma_i(t_k) + \frac{\|g\|}{\|g_i\|}[\gamma_1(k+2,\lambda) - \gamma_1(k+1,\lambda)] \\
&= \gamma_i(t_k) + \frac{\|g_i(\lambda)\|}{\|g_i\|}[\gamma_i(k+2,\lambda) - \gamma_i(k+1,\lambda)] \\
&< \gamma_i(t_k) + [\gamma_i(k+2,\lambda) - \gamma_i(k+1,\lambda)] \\
&< \gamma_i(k+2,\lambda).
\end{aligned} \tag{73}$$

The above inequality and the fact that $\gamma_i(t_{k+1}) = 0$ for all $i > M$ imply that

$$\gamma_i(t_{k+1}) < \gamma_i(k+2,\lambda) \tag{74}$$

for all $i = 2, ..., k$.

If $t_{k+1} \le k+1$ then $\gamma_{k+1}(t_{k+1}) = 0$ which shows that

$$\gamma_{k+1}(t_{k+1}) < \gamma_{k+1}(k+2,\lambda). \tag{75}$$

Otherwise, we have that $t_{k+1} \in (k+1, k+2)$.

$$\begin{aligned}\gamma_{k+1}(t_{k+1}) &= \gamma_{k+1}(t_{k+1}) - \gamma_{k+1}(k+1) = \frac{\|g\|}{\|g_{k+1}\|}\left[\gamma_1(t_{k+1}) - \gamma_1(k+1)\right] \\ &< \frac{\|g\|}{\|g_{k+1}\|}\left[\gamma_1(t_{k+1}) - \gamma_1(t_k)\right] = \frac{\|g\|}{\|g_{k+1}\|}\left[\gamma_1(k+2,\lambda) - \gamma_1(k+1,\lambda)\right] \\ &= \frac{\|g\|}{\|g_{k+1}\|}\frac{\|g_{k+1}(\lambda)\|}{\|g(\lambda)\|}\left[\gamma_{k+1}(k+2,\lambda) - \gamma_{k+1}(k+1,\lambda)\right] \\ &< \gamma_{k+1}(k+2,\lambda) - \gamma_{k+1}(k+1,\lambda) = \gamma_{k+1}(k+2,\lambda). \end{aligned} \tag{76}$$

The inequalities (74)–(76) show that (63) holds when k is replaced by $k+1$. By induction, we see that the theorem is true.

□

Lemma 7. *For any $\lambda > 0$ and $k \geq 1$ such that $g_k \neq 0$, there exists a unique $\hat{t} \in [1, k+1)$ such that*

$$\|x(\hat{t})\| = \|x_{k+1}(\lambda)\|, \tag{77}$$

furthermore

$$-x_{k+1}(\lambda)^T g \leq -x(\hat{t})^T g. \tag{78}$$

If $k > 1$ the above inequality holds as a strictly inequality.

Proof. From the above theorem, there exists a $t_k \in [1, k+1)$ such that (62) and (63) hold. These two relations gives that

$$\|x_{k+1}(\lambda)\| \geq \|x(t_k)\|. \tag{79}$$

The above inequality holds as a strictly inequality if $k > 1$. The monotone property of $\|x(t)\|$ and inequality (79) indicates that there exists a $\hat{t} \in [t_k, k+1)$ such that

$$\|x(\hat{t})\| = \|x_{k+1}(\lambda)\|. \tag{80}$$

Because $\hat{t} \geq t_k$, we have that

$$-x(\hat{t})^T g \geq -x(t_k)^T g = -x_{k+1}(\lambda)^T g. \tag{81}$$

If $k > 1$, (79) holds as a strictly inequality, which implies that $\hat{t} > t_k$, consequently we see that (81) holds as a strictly inequality.

□

Theorem 2. *For any $\Delta > 0$. $g \in \Re^n$ and any positive definite matrix $B \in \Re^{n\times n}$, let d^* be the global solution of the trust region subproblem (2)–(3), and let x^* be the solution obtained by the truncated CG method, then inequality (18) holds.*

Proof. Due to the fact that $\phi(0) = 0$, (18) is equivalent to

$$\phi(x^*) \leq \frac{1}{2}\phi(d^*). \tag{82}$$

Thus we only need to prove the above inequality. Because d^* is the exact solution of (2)–(3), it follows from Lemma 1 that there exists a Lagrange multiplier $\lambda \geq 0$ such that d^* is the minimizer of $\phi(d) + \frac{1}{2}\lambda\|d\|^2$. If $\lambda = 0$, then $x^* = d^*$, which implies (82).

Now we assume that $\lambda > 0$. There exists a $k \in [1, n]$ such that $g_k(\lambda) \neq 0$ and $g_{k+1}(\lambda) = 0$. This is easy to see that

$$d^* = x_{k+1}(\lambda). \tag{83}$$

From the above lemma, there exists a $\hat{t}$ such that (77) and (78) hold. $\lambda > 0$ implies that $\|d^*\| = \Delta$, Thus, (83) and (77) show that $\|x(\hat{t})\| = \Delta$. Therefore, by the definition of the CG path $x(t)$ we have that

$$x^* = x(\hat{t}). \tag{84}$$

Thus, it follows from (48) and (78) that

$$\phi(x^*) = \frac{1}{2}g^T x^* + \frac{1}{2}x(\hat{t})^T g(x(\hat{t})) \leq \frac{1}{2}g^T x^* \leq \frac{1}{2}g^T x_{k+1}(\lambda) = \frac{1}{2}g^T d^* \leq \frac{1}{2}\phi(d^*). \tag{85}$$

This completes our proof.

□

We have shown that if the Steihaug-Toint truncated CG method is used to solve the trust region subproblem (2)–(3), the function reduction is at least half of the maximum reduction.

Preconditions change the variable d without alternating the function value $\phi(d)$, therefore it is easy to see our result is independent of preconditioning.

Acknowledgements. The author would like to thank Professor M.J.D. Powell and Professor Ph. Toint for their comments on an earlier version of the paper.

References

1. Chen, X. (1998): Trust Region Subproblems and Extensions, in Chinese. Master Thesis, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing
2. Fletcher, R. (1987): Practical Methods of Optimization, second edition. John Wiley and Sons, Chichester
3. Gay, D.M. (1981): Computing optimal local constrained step. SIAM J. Sci. Stat. Comp. **2**, 186–197
4. Moré, J.J., Sorensen, D.C. (1983): Computing a trust region step. SIAM J. Sci. Stat. Comp. **4**, 553–572
5. Powell, M.J.D. (1975): Convergence properties of a class of minimization algorithms. In: Mangasarian, O.L., Meyer, R.R., Robinson, S.M., eds., Nonlinear Programming 2, pp. 1–27. Acadmic Press, New York
6. Steihaug, T. (1983): The conjugate gradient method and trust regions inlarge scale optimization. SIAM J. Numer. Anal. **20**, 626–637
7. Toint, Ph.L. (1981): Towards an efficient sparsity exploiting Newton method for minimization. In: Duff, I., ed., Sparse Matrices and Their Uses, pp. 57–88. Academic Press
8. Tseng, P. (1998): Private communications
9. Ye, Y. (1997): Approximating quadratic programming with bound and quadratic constraints. Report, Dept. Management Sci., University of Iowa, Iowa 52242, USA
10. Yuan, Y. (1997): Some properties of a trust region subproblem. Present in the 16th International Symposium on Mathematical Programming, Lausanne, Switzerland, August 24–29