ANALYSIS ON THE CONJUGATE GRADIENT METHOD

YA-XIANG YUAN

Computing Center, Academia Sinica, Beijing 100080, China

(Received 12 October 1992; in final form 17 January 1993)

In this paper we analyze the conjugate gradient method when the objective function is quadratic. We apply backward analyses to study the quadratic termination of the conjugate gradient method. Forward analyses are used to derive some properties of the conjugate gradient method, including the only linear convergence of the method and an upper bound for the rate of convergence.

KEY WORDS: Conjugate gradient, quadratic termination, linear convergence.

1 INTRODUCTION

Conjugate gradient method is one of the basic numerical methods for the unconstrained optimization problem:

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

where $f(x)$ is a nonlinear function defined in R^n. Conjugate gradient algorithms are iterative. The initial point x_1 is given and the first search direction d_1 is chosen by

$$d_1 = -g_1 = -g(x_1) = -\nabla f(x_1). \tag{2}$$

We use the notations $g_k = g(x_k) = \nabla f(x_k)$. At the k-th iteration, given the current iterate x_k and the search direction d_k, a step-length α_k is calculated by the exact line search:

$$\alpha_k = arg\ \min\{f(x_k+\alpha d_k); \quad \alpha > 0\} \tag{3}$$

and the next iterate is set to

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

The search direction for the next iteration is defined by

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

such that d_{k+1} is "conjugate to" d_k in the sense that

$$d_{k+1}^T A d_k = 0 \tag{4}$$

if the objective function is a convex quadratic function

$$f(x) = \frac{1}{2}x^T Ax + b^T x + c \tag{5}$$

where A is an $n \times n$ positive definite matrix. The exact line search (3) implies that

$$g_{k+1}^T d_k = 0. \tag{6}$$

More details about the conjugate gradient method can be found in [1], [3] and [8].

Assume that $f(x)$ is (5), in order to force (4), it follows that

$$\beta_k = \frac{g_{k+1}^T A d_k}{d_k^T A d_k}. \tag{7}$$

Because $f(x)$ is convex quadratic, it can be shown that (7) is the same as

$$\beta_k = \frac{\|g_{k+1}\|_2^2}{\|g_k\|_2^2} \tag{8}$$

or

$$\beta_k = \frac{g_{k+1}^T (g_{k+1} - g_k)}{\|g_k\|_2^2}, \tag{9}$$

if the first search direction d_1 is chosen by (2). (8) and (9) give the Fletcher-Reeves method [4] and the Polak-Rabiére-Polyak method [5], [6] respectively.

One nice property of the conjugate gradient method is that it terminates after at most n iterations if $f(x)$ is a convex quadratic function and if the first search direction is chosen by (2). However, for a general nonlinear function $f(x)$, the objective function $f(x)$ can be closely approximated by a quadratic function only after certain number of iterations. Thus local analysis can not apply to show quadratic termination because it is usual that $d_k \neq -g_k$ for $k > 1$, due to early iterations. Hence it is important to study the performance of the conjugate gradient method when the condition (2) is violated. Instead, we only assume that $d_1^T g_1 < 0$.

It was first shown by Powell [7] that the conjugate gradient method normally does not terminate after finite many iterations if the condition (2) is removed, though the objective function $f(x)$ is quadratic and strictly convex. Powell [7] also pointed out if termination happens it must be within the first $n+1$ iterations. We re-prove this result by using backward analyses. It is known that, the conjugate gradient method may converge similar to that of the steepest descent direction method if condition (2) is not satisfied (see, [2]). We give a theoretical study, and show that the conjugate gradient method applied to convex quadratical functions converges always only linearly if finite termination does not happen. An upper bound for the linear rate of convergence is also given.

As presented in [7], throughout the rest of this paper, we assume that the objective function is the simple convex quadratic function:

$$f(x) = \frac{1}{2}x^T Ax$$

where A is an $n \times n$ symmetric positive definite matrix.

In the next section, we introduce backward analyses to derive some results on quadratic termination, including the main results in [7]. In section 3, we use forward analyses to show that the rate of conjugate gradient method is only linear if it does not terminate within the first $n+1$ iterations.

2 BACKWARD ANALYSIS

In this section, we analyze the conjugate gradient method by backward analyses. Under the assumption that termination happens at the N-th iteration, we prove that N must not be greater than $n+1$.

First we recall the following relations:

$$x_{k+1} = x_k + \alpha_k d_k$$
$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{10}$$

and

$$\alpha_k = \frac{-g_k^T d_k}{d_k^T A d_k}, \qquad \beta_k = \frac{g_{k+1}^T A d_k}{d_k^T A d_k} \tag{11}$$

which imply that

$$g_{k+1} = g_k + \alpha_k A d_k \tag{12}$$
$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|_2^2$$

for all k.

The following Lemmas 2.1-2.3 are straightforward:

Lemma 2.1. *For all $k \geq 2$, we have*

$$d_k^T A(d_k + g_k) = 0 \tag{13}$$
$$g_k^T (d_k + g_k) = 0. \tag{14}$$

Proof. The lemma is obvious if $d_k + g_k = 0$. Assume that $d_k + g_k \neq 0$, it can be seen that $\beta_{k-1} \neq 0$. Therefore we have that

$$d_{k-1} = \frac{1}{\beta_{k-1}}(d_k + g_k).$$

Now equation (13) follows from the conjugate condition $d_{k-1}^T A d_k = 0$ and (14) from the exact line search condition (6). □

Lemma 2.2. *If equation*

$$d_k + g_k = 0 \tag{15}$$

is true, then either $g_k = d_k = 0$ or $k \leq 2$.

Proof. Due to equation (15) and the conjugate condition (4) we have that

$$d_{k-1}^T A g_k = -d_{k-1}^T A d_k = 0. \tag{16}$$

Now relations (16) and (12) imply that

$$d_{k-1}^T A g_{k-1} - g_{k-1}^T d_{k-1} \frac{\|Ad_{k-1}\|_2^2}{d_{k-1}^T A d_{k-1}} = 0. \tag{17}$$

Assume that $k > 2$, Lemma 2.1 and the fact that $k - 1 \geq 2$ show that

$$g_{k-1}^T d_{k-1} = -\|g_{k-1}\|_2^2 \tag{18}$$

$$d_{k-1}^T A d_{k-1} = -d_{k-1}^T A g_{k-1}. \tag{19}$$

Substituting (18) and (19) into equation (17), we find that

$$(d_{k-1}^T A g_{k-1})^2 = \|g_{k-1}\|_2^2 \|Ad_{k-1}\|_2^2$$

which proves that the vectors g_{k-1} and Ad_{k-1} are collinear. Consequently $g_k = 0$. □

Lemma 2.3. *The relation $g_{k+1}^T g_k = 0$ holds for all $k \geq 2$.*

Proof. By the conjugate condition (4) and the exact line search condition (6) we have that $d_{k-1}^T g_{k+1} = d_{k-1}^T g_k + \alpha_k d_{k-1}^T A d_k = 0$. Now by (10) the relation $g_{k+1}^T g_k = g_{k+1}^T(-d_k + \beta_{k-1} d_{k-1}) = 0$ holds for all $k \geq 2$. □

For the rest of this section, we assume that the conjugate gradient method terminates after N iterations, that is

$$g_{N+1} = 0. \tag{20}$$

Termination condition (20) indicates that the vectors g_N and Ad_N are collinear. Thus there exists a number ρ such that $d_N = -\rho A^{-1} g_N$.

Under condition (20), we have the following results:

Lemma 2.4. *If*

$$d_k + g_k \neq 0 \tag{21}$$

for $k = N, N-1, ..., N-l+1$, then

$$d_i^T A d_j = 0 \tag{22}$$

$$g_i^T g_j = 0 \tag{23}$$

for $N - l \leq i < j \leq N$.

Proof. The lemma is obviously true if $l = 1$. Now we assume that it is true for $l = l_0$ and prove the lemma for $l = l_0 + 1$. It is easily seen that we only need to prove equations (22) and (23) for $i = N - l_0 - 1$ and $i < j \leq N$.

$$\begin{aligned} d_{N-l_0-1}^T A d_j &= \frac{1}{\beta_{N-l_0-1}} (d_{N-l_0} + g_{N-l_0})^T A d_j = \frac{1}{\beta_{N-l_0-1}} g_{N-l_0}^T A d_j \\ &= \begin{cases} -\frac{\rho}{\beta_{N-l_0-1}} g_{N-l_0}^T g_N, & \text{if } j = N \\ \frac{1}{\alpha_j \beta_{N-l_0-1}} g_{N-l_0}^T (g_{j+1} - g_j), & \text{if } j < N \end{cases} = 0 \end{aligned}$$

and

$$\begin{aligned} g_{N-l_0-1}^T g_j &= (g_{N-l_0} - \alpha_{N-l_0-1} A d_{N-l_0-1})^T g_j \\ &= -\alpha_{N-l_0-1} d_{N-l_0-1}^T A g_j \\ &= -\alpha_{N-l_0-1} d_{N-l_0-1}^T A (d_j - \beta_{j-1} d_{j-1}) = 0 \end{aligned}$$

Therefore the lemma follows by induction. ∎

Lemma 2.5. *Assume the conjugate gradient method terminates after N iterations, then* $N \leq n+1$.

Proof. Suppose that $N > n+1$, Lemma 2.2 implies that (21) holds for $l = N-2$. Thus we have that

$$g_i^T g_j = 0 \tag{24}$$

for $2 \leq i < j \leq N$. On the other hand, our assumption implies that $g_i (i = 2, 3, ..., N)$ are all non-zero vectors. Now (24) is impossible because there do not exist $N-1 (> n)$ non-zero mutual conjugate orthogonal vectors in R^n. This contradiction shows that $N \leq n+1$. ∎

Let $\tilde{d}_k$ and $\tilde{g}_k$ be defined as follows:

$$\tilde{d}_N = \rho d_N, \qquad \tilde{d}_k = \rho \prod_{i=k}^{N-1} \beta_i d_k \tag{25}$$

$$\tilde{g}_N = g_N, \qquad \tilde{g}_k = \rho \alpha_k^{-1} \prod_{i=k}^{N-1} \beta_i g_k \tag{26}$$

for $k < N$. From (10), (12) and (25)-(26), we have that

$$\tilde{d}_{k-1} = \tilde{d}_k + \rho^{-1} \alpha_k \tilde{g}_k \tag{27}$$

$$\tilde{g}_{k-1} = \alpha_k \alpha_{k-1}^{-1} \beta_k^{-1} \tilde{g}_k - A \tilde{d}_{k-1}. \tag{28}$$

Set $\bar{d}_k = A^{1/2} \tilde{d}_k$, $\bar{g}_k = A^{-1/2} \tilde{g}_k$, $\bar{\beta}_k = \rho^{-1} \alpha_k$, $\bar{\alpha}_k = \alpha_k \alpha_{k-1}^{-1} \beta_k^{-1}$. It follows from the relations (27)-(28) that

$$\bar{d}_{k-1} = \bar{d}_k + \bar{\beta}_k A \bar{g}_k \tag{29}$$

$$\bar{g}_{k-1} = -\bar{d}_{k-1} + \bar{\alpha}_k \bar{g}_k, \tag{30}$$

for $k \leq N$. Notice that

$$\bar{g}_N = -\bar{d}_N. \tag{31}$$

Now we can easily see that (29)-(31) are exactly conjugate gradient iterations except they are backward, as the "conjugate condition" $\bar{g}_{k-1}^T A \bar{g}_k = 0$ and the "exact line search condition" $\bar{d}_{k-1}^T \bar{g}_k = 0$ are satisfied. Hence we can trace the iterations backward provided that (15) does not hold.

Assume that $N > n+1$, the backward iterations (29)-(31) and the standard analysis of the conjugate gradient method imply that $\bar{g}_{N-n} = 0$, which yields $d_{N+1-n} + g_{N+1-n} = 0$, but it contradicts Lemma 2.2 because $N + 1 - n > 2$. This contradiction verifies the validity of Lemma 2.5.

3 FORWARD ANALYSIS

It is known that termination happens if $d_1 + g_1 = 0$ or $d_2 + g_2 = 0$ (for example, see [3]). Without loss of generality, due to Lemma 2.2, we assume that,

$$d_k + g_k \neq 0 \tag{32}$$

for all k.

First we have the following results:

Lemma 3.1. *For $k > 1$, we have*

$$Ad_k = -\frac{1}{\alpha_k}[d_{k+1} - (1+\beta_k)d_k + \beta_{k-1}d_{k-1}] \tag{33}$$

where α_k and β_k are defined by (11).

Proof. Equations (12) and (10) can be rewritten as

$$Ad_k = \frac{1}{\alpha_k}(g_{k+1} - g_k) \tag{34}$$

$$g_{k+1} = -d_{k+1} + \beta_k d_k. \tag{35}$$

Since $k > 1$, we can replace k by $k-1$ in equation (35), which gives that

$$g_k = -d_k + \beta_{k-1}d_{k-1}. \tag{36}$$

Now (33) follows from (34), (35) and (36). □

Lemma 3.2. *For any integer $l \geq 1$, if $d_k, d_{k+1}, ..., d_{k+l}$ are mutual conjugate, then*

$$d_j^T g_i = 0\,, \qquad k \leq j < i \leq k+l+1. \tag{37}$$

Proof. From (12), (6) and the assumption of the lemma, it follows that

$$d_j^T g_i = d_j^T[g_{j+1} + \sum_{t=j+1}^{i-1} \alpha_t A d_t] = 0$$

which verifies (37). □

Lemma 3.3. *If $d_k, d_{k+1}, ..., d_{k+l}$ are mutual conjugate, then the vectors d_{k+1}, $d_{k+2}, ..., d_{k+l+1}$ are also mutual conjugate.*

Proof. It is sufficient to prove that

$$d_{k+l+1}^T A d_j = 0 \tag{38}$$

for $j = k+1, ..., k+l$. As (38) is obvious if $j = k+l$, we only need to show (38) for $k < j < k+l$. Applying (10), (33) and (37), we can show that

$$\begin{aligned} d_{k+l+1}^T A d_j &= (-g_{k+l+1} + \beta_{k+l}d_{k+l})^T A d_j = -g_{k+l+1}^T A d_j \\ &= \frac{1}{\alpha_j} g_{k+l+1}^T (d_{j+1} - (1+\beta_j)d_j + \beta_{j-1}d_{j-1}) = 0, \end{aligned}$$

which gives relation (38) and consequently the lemma is true. ∎

Let L be the largest non-negative integer such that $\{d_1, d_2, ..., d_{L+1}\}$ are non-zero mutual conjugate. ¿From Lemma 3.3, it is clear that either

$$d_{L+2} = 0 \tag{39}$$

or

$$d_{L+2}^T A d_1 \neq 0. \tag{40}$$

Since (39) implies $g_{L+2} = 0$ which indicates that termination happens at the $(L+1)$-th iteration, we only consider the case that (40) holds. (40) and (4) imply that $L \geq 1$. It is easy to show the following lemma.

Lemma 3.4. *If $d_1, ..., d_{L+1}$ are mutual conjugate and (40) holds, then we have that*

$$d_{L+3}^T A d_2 \neq 0 \tag{41}$$

Proof. Due to Lemma 3.1 and Lemma 3.2, it follows that

$$\begin{aligned}
d_{L+3}^T A d_2 &= (-g_{L+3} + \beta_{L+2} d_{L+2})^T A d_2 \\
&= -g_{L+3}^T A d_2 \\
&= \frac{1}{\alpha_2} g_{L+3}^T (d_3 - (1+\beta_2) d_2 + \beta_1 d_1) \\
&= \frac{\beta_1}{\alpha_2} g_{L+3}^T d_1 \\
&= \frac{\beta_1}{\alpha_2} (g_{L+2} + \alpha_{L+2} A d_{L+2})^T d_1 \\
&= \frac{\beta_1}{\alpha_2} \alpha_{L+2} d_{L+2}^T A d_1 \neq 0,
\end{aligned} \tag{42}$$

where the last part of (42) follows from (41) and that (32) implies $\beta_1 \neq 0$. ∎

Consequently, we have the following corollary:

Corollary 3.5. *Under the assumptions of Lemma 3.4, $d_k, d_{k+1}, ..., d_{k+L}$ are mutual conjugate, but $d_{k+L+1}^T A d_k \neq 0$ for all k.*

The following lemma indicates that the ratio $\|d_k\|_2/\|g_k\|_2$ is bounded below and above uniformly for all k.

Lemma 3.6. *For all $k > 1$, the following inequalities hold:*

$$1 \leq \|d_k\|_2/\|g_k\|_2 \leq \sqrt{k_2(A)}, \tag{43}$$

where $k_2(A) = \lambda_1(A)/\lambda_n(A)$. $\lambda_1(A)$ and $\lambda_n(A)$ are the largest and the smallest eigenvalues of A.

Proof. The first part of (43) is trivial due to (10) and (6). It follows from (13) that $(d_k^T A d_k)^2 = (d_k^T A g_k)^2 \leq ||Ad_k||_2^2 ||g_k||_2^2$, which yields

$$\frac{\|d_k\|_2^2}{\|g_k\|_2^2} \leq \frac{\|d_k\|_2^2}{\|g_k\|_2^2} \frac{\|g_k\|_2^2 \|Ad_k\|_2^2}{(d_k^T A d_k)^2} = \frac{\|d_k\|_2^2 \|Ad_k\|_2^2}{(d_k^T A d_k)^2} \leq k_2(A). \tag{44}$$

Now, (44) gives the second part of (43). ∎

Lemma 3.7. *For all $k > 1$, we have that*

$$\frac{1}{k_2(A)\lambda_1(A)} \le \alpha_k \le \frac{1}{\lambda_n(A)}.$$

Proof. It follows from (11), (13) and (43) that

$$\alpha_k = \frac{-g_k^T d_k}{d_k^T A d_k} = \frac{\|g_k\|_2^2}{d_k^T A d_k} \le \frac{\|d_k\|_2^2}{d_k^T A d_k} \le \frac{1}{\lambda_n(A)}$$

and that

$$\alpha_k = \frac{\|g_k\|_2^2}{d_k^T A d_k} \ge \frac{\|d_k\|_2^2}{k_2(A) d_k^T A d_k} \ge \frac{1}{k_2(A)\lambda_1(A)}.$$

∎

It is known that the conjugate gradient method converges at least linearly, as at each iteration it improves the objective function value as good as the steepest descent method. That is, we can show the following inequality

$$\frac{f(x_{k+1}) - f(x^*)}{f(x_k) - f(x^*)} = \frac{g_{k+1}^T A^{-1} g_{k+1}}{g_k^T A^{-1} g_k} = 1 - \frac{(g_k^T d_k)^2}{d_k^T A d_k g_k^T A^{-1} g_k} \le 1 - \frac{1}{k_2(A)},$$

where $x^* = 0 \in R^n$ is the solution of (1).

However, if the conjugate gradient method does not terminate within the first $n+1$ iterations, its convergence rate is only linear. We shall give an upper bound for the rate of convergence of the method. The following two lemmas are needed to obtain the upper bound.

Lemma 3.8. *Under the conditions of Lemma 3.4, we have that $g_{L+3}^T d_1 \ne 0$ and that the recurrence relation*

$$g_{L+k+3}^T d_{k+1} = \frac{\alpha_{L+k+2}}{\alpha_{k+1}} \beta_k g_{L+k+2}^T d_k \tag{45}$$

holds for all k.

Proof. From (42), it is easy to see that $g_{L+3}^T d_1 \ne 0$. Similar to (42), we can show that

$$g_{L+k+3}^T d_{k+1} = \alpha_{L+k+2} d_{L+k+2}^T A d_{k+1} = \alpha_{L+k+2} \frac{\beta_k}{\alpha_{k+1}} g_{L+k+2}^T d_k,$$

which gives (45). ∎

Lemma 3.9. *Without the condition (2), the definitions (7), (8) and (9) are equivalent for all $k > 1$.*

Proof. Assume $k > 1$, it follows from (33), (10) and (18) that

$$\frac{g_{k+1}^T A d_k}{d_k^T A d_k} = -\frac{g_{k+1}^T [d_{k+1} - (1+\beta_k) d_k + \beta_{k-1} d_{k-1}]}{\alpha_k d_k^T A d_k} = \frac{\|g_{k+1}\|_2^2}{\|g_k\|_2^2}, \tag{46}$$

which shows that (7) and (8) are equivalent. Consequently the lemma follows from (46) and Lemma 2.3. ∎

Now we can prove the following result on the only linear convergence of the conjugate gradient method, which also provides an upper bound for the rate of convergence.

Theorem 3.10. *Assume (40) holds, then the conjugate gradient method converges only linearly. Furthermore, the following inequality*

$$\|g_{k+L+2}\|_2 \geq \frac{|\beta_1 g_{L+3}^T d_1|}{\|g_2\|_2^2 [k_2(A)]^{2L+2.5}} \|g_k\|_2 \tag{47}$$

holds for all k.

Proof. From Lemmas 3.7 and 3.8, we can derive that

$$\begin{aligned} |g_{k+L+2}^T d_k| &= |d_1^T g_{L+3}| \prod_{i=1}^{k-1} \frac{\alpha_{L+i+2}|\beta_i|}{\alpha_{i+1}} \\ &= |d_1^T g_{L+3}\beta_1| \frac{\|g_k\|_2^2}{\|g_2\|_2^2} \prod_{i=1}^{L+1} \frac{\alpha_{k+i}}{\alpha_{i+1}} \\ &\geq |d_1^T g_{L+3}\beta_1| \frac{\|g_k\|_2^2}{\|g_2\|_2^2} \left[\frac{1}{k_2(A)}\right]^{2(L+1)} . \end{aligned} \tag{48}$$

On the other hand,

$$|g_{k+L+2}^T d_k| \leq \|g_{k+L+2}\|_2 \|d_k\|_2 \leq \|g_{k+L+2}\|_2 \|g_k\|_2 \sqrt{k_2(A)}. \tag{49}$$

Now inequality (47) follows from (48) and (49). As $g_{L+3}^T d_1 \neq 0$, and $d_2 \neq g_2$ implies that $\beta_1 \neq 0$, it follows from (47) that the conjugate gradient method converges only linearly. ∎

Inequality (47) shows that the conjugate gradient method converges always only linearly if finite termination does not happen. It should be mentioned that (47) is only an upper bound for the convergence, which does not implies that the conjugate gradient method converges always with the rate of this upper bound. It is also possible that this upper bound can be reduced. This seems reasonable, because some examples are known where, the convergence of the conjugate gradient method is indeed the same as in the steepest descent method (see, [2]).

REFERENCES

1. A. Buckley, *Conjugate gradient methods*, in: Nonlinear Optimization 1981, M.J.D. Powell, ed., Academic Press, London, 1982, pp. 17–22.
2. H.P.Crowder and P.Wolfe, *Linear convergence of the conjugate gradient method*, IBM J. Res. Develop., 16 (1972), pp. 431–433.
3. R. Fletcher, *Practical Optimization: Vol 1, Unconstrained Optimization*, John Wiley and Sons, Chichester, 1980.
4. R. Fletcher and C.M. Reeves, *Function minimization by conjugate gradients*, Computer Journal 7 (1964), pp. 149–154.
5. E. Polak, and G. Ribiére, *Note sur la convergence de méthodes de directions conjugees*, Rev. Fr. Inform. Rech. Oper. 16 (1969), pp. 35–43.

6. B.T. Polyak, *The conjugate gradient method in extremal problems*, USSR Comp. Maths. and Math. Phys., 9 (1969), pp. 94–112.
7. M.J.D. Powell, *Some convergence properties of the conjugate gradient method*, Math. Prog., 11 (1976), pp. 42–49.
8. M.J.D. Powell, *Nonconvex minimization calculations and the conjugate gradient method*, in: Numerical Analysis, G.A. Watson, ed., Lecture Notes in Mathematics, Vol. 1066, Springer–Verlag, London, 1984, pp. 121–141.