A Variable Metric Algorithm with Broyden Rank One Modifications for Nonlinear Equality Constraints Optimization

Chunyan Hu1, Zhibin Zhu2

1School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin, China
2School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, China
Email: huchyel@hotmail.com

Received January 17, 2013; revised February 18, 2013; accepted March 8, 2013

ABSTRACT

In this paper, a variable metric algorithm is proposed with Broyden rank one modifications for the equality constrained optimization. This method is viewed as an expansion in constrained optimization as the quasi-Newton method to unconstrained optimization. The theoretical analysis shows that local convergence can be induced under some suitable conditions. In the end, it is established an equivalent condition of superlinear convergence.

Keywords: Equality Constrained Optimization; Variable Metric Algorithm; Broyden Rank One Modification; Superlinear Convergence

1. Introduction & Algorithm

This paper proposes to consider the following nonlinear mathematical programming problem:

\[ \min f(x) \quad \text{s.t.} \quad g_j(x) = 0, \quad j \in I = \{1, \ldots, m\}, \]

where \( f: \mathbb{R}^n \to \mathbb{R} \) and \( g_j: \mathbb{R}^n \to \mathbb{R} \) are continuously differentiable functions. Denote the feasible set as follows:

\[ X = \{ x \in \mathbb{R}^n \mid g_j(x) = 0, \quad j \in I \}. \]

Let \( L(x, \lambda) \) be the Lagrangian function of (1), and

\[ L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x). \]

If \((x_k, \lambda_k)\) is a KKT point pair of Equation (1), then

\[ \nabla L(x_k, \lambda_k) \bigg|_{x=x_k, \lambda=\lambda_k} = 0, \quad \text{i.e.,} \]

\[ \nabla f(x_k) + \nabla g(x_k) \lambda_k = 0, \]

\[ g(x_k) = 0, \]

where \( g(x_k) = \{ g(x_k), j \in I \} \). At the point pair \((x_k, \lambda_k)\), the Newton’s iteration of (2) is defined as follows:

\[ \begin{bmatrix} \nabla^2 L(x_k, \lambda_k) & \nabla g(x_k) \\ \nabla^T g(x_k) & 0 \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix} = - \begin{bmatrix} \nabla f(x_k) + \nabla g(x_k) \lambda_k \\ g(x_k) \end{bmatrix}. \]

Later, a positive definite matrix \( H_k^{-1} \) is replaced for \( \nabla^2 L(x_k, \lambda_k) \) by a lot of authors to develop some kinds of variable metric methods, such as sequential quadratic programming (SQP) methods [1-5], sequential systems of linear equations (SSLE) algorithms [6-8]. In general, the computational cost of those methods is large.

In this paper, a new variable metric method is presented, in which the following fact is based on: a positive definite matrix \( B_k \) is replaced for the matrix

\[ \begin{bmatrix} \nabla^2 L(x_k, \lambda_k) & \nabla g(x_k) \\ \nabla^T g(x_k) & 0 \end{bmatrix}. \]

In the sequel, we describe the algorithm for the solution of (1). Denote

\[ z = \begin{bmatrix} x \\ \lambda \end{bmatrix}, \quad z_k = \begin{bmatrix} x_k \\ \lambda_k \end{bmatrix}, \quad B_k = \begin{bmatrix} \nabla^2 L(x_k, \lambda_k) & \nabla g(x_k) \\ \nabla^T g(x_k) & 0 \end{bmatrix}, \]

\[ v(z_k) = -\begin{bmatrix} \nabla f(x_k) + \nabla g(x_k) \lambda_k \\ g(x_k) \end{bmatrix}, \quad p_k = z_{k+1} - z_k = \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix}, \]

\[ q_k = v(z_k), \quad F(z) = \| \nabla f(x) + \nabla g(x) \lambda \|^2 + \| g(x) \|^2. \]

It is obvious that \( \min_{x \in X} f(x) \Leftrightarrow \min_{z \in \mathbb{R}^{m+n}} F(z) \), and from Equation (3), we have

\[ B_k p_k = v(z_k). \]
To the system of linear Equations (4), like unconstrained optimization, $B_k$ is dealt with using Broyden rank one modifications as follows:

$$B_{k+1} = B_k + \frac{(q_k - B_k p_k) p_k^T}{p_k^T p_k}.$$  \hfill (5)

Now, the algorithm for the solution of Equation (1) can be stated as follows:

**Algorithm A:**

Step 1: Initialization: Given a starting point $z_0 \in \mathbb{R}^{n+m}$ (i.e., $x_0 \in \mathbb{R}^n, \lambda_0 \in \mathbb{R}^m$), and a initial positive definite matrix $B_0 \in \mathbb{R}^{(n+m)\times(n+m)}$. $\varepsilon > 0, k = 0$;

Step 2: Compute $v(z_k)$. If $\|v(z_k)\| \leq \varepsilon$, stop;

Step 3: Compute $d_k = -B_k^{-1}v(z_k)$;

Step 4: Let $z_{k+1} = z_k + d_k$, and obtain $B_{k+1}$ according to (5). Set $k = k + 1$. Go back to step 2.

2. Convergence of Algorithm

If the algorithm stops at $z_k = \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix}$, then $x_k$ is a KKT point of (1). In the sequel, we suppose that algorithm generates an infinite sequence $\{z_k\}$.

Four basic assumptions are given as follows:

A1 The feasible set $X$ is nonempty; The functions $f, g_j$ ($j \in I$) are two-times continuously differentiable;

A2 For all $x \in \mathbb{R}^n$, the vectors $\{\nabla g_j(x), j \in I\}$ are linearly independent;

A3 $\{x_k\}$ and $\{\lambda_k\}$ are bounded. There exists a KKT point pair $(x_k, \lambda_k)$, such that $\nabla^2 L(x_k, \lambda_k)$ is positive definite;

A4 There exists a ball $N(x_k, \delta)$ of radius $\delta > 0$ about $x_k$, where $\nabla^2 f(x), \nabla^2 g_j(x), \nabla g_j(x) (j \in I)$ satisfy the Lipschitz condition on $N(x_k, \delta)$.

**Lemma 1** [9] Let $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$ be continuously differentiable in some open and convex set $D$, and $F'$ is Lipschitz continuous in $D$, then $\forall x + d \in D$, it holds that

$$\|F(x + d) - F(x) - F'(x)d\| \leq \frac{\nu}{2}\|d\|^2,$$

where $\nu$ is the Lipschitz constant. Moreover, $\forall u, v, x \in D$, it follows that

$$\|F(u) - F(v) - F'(u - v)v\| \leq \nu\|u - v\|\|v - x\|/2 + \|u - v\|\|u - v\|.$$

**Lemma 2** [9] Let $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$ be continuously differentiable in some open and convex set $D$, and $F'$ is Lipschitz continuous in $D$. Moreover, assume $F'(x)$ is invertible for some $x \in D$, then there exist some $\varepsilon > 0, \beta > \alpha > 0$, such that for all $u, v \in D$, the fact $\{\|u - x\|, \|v - x\|\} \leq \varepsilon$ implies that

$$\alpha\|u - v\| \leq \|F(u) - F(v)\| \leq \beta\|u - v\|.$$

**Lemma 3** [9] For operator $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m}$, which satisfies:

R1 $F$ is continuously differentiable on $D$;

R2 There exists a point $z_0 \in D$, such that $F(z_0) = 0$, and $F'(z_0)$ is reversible;

R3 $F'$ is Lipschitz continuous at $z_0$, i.e., there exists a constant $\nu$, such that

$$\|F'(z) - F'(z_0)\| \leq \nu\|z - z_0\|$$

for all $z \in D$.

Let $z_{k+1} = z_k - H_k^{-1}F(z_k)$, where $H_k^{-1} \in \mathbb{R}^{(n+m)\times(n+m)}$, and it holds that

$$\left\|H_k - F'(z_k)\right\| \leq \frac{\nu}{2}\left(\|z_{k+1} - z_k\| + \|z_k - z_0\|\right), \forall k,$$

then there exist $\varepsilon > 0$ and $\delta > 0$, when

$$\|z_0 - z_k\| < \varepsilon, \|H_k - F'(z_k)\| < \delta,$$

it is true that $z_{k+1}$ is meaning, and $z_k$ is linearly convergent to $z_0$. Therefore, we can conclude that $z_k$ is superlinearly convergent to $z_0$, if and only if

$$\lim_{k \rightarrow \infty} \frac{\|H_k - F'(z_k)\|}{\|z_{k+1} - z_k\|} = 0.$$ \hfill (7)

In the sequel, we prove the convergence Theorem as follows:

**Theorem 1** If there exist constants $\varepsilon > 0$ and $\delta > 0$, such that

$$\|z_0 - z_k\| < \varepsilon, \|B_0 - v'(z_k)\| < \delta,$$

then $\{z_k\}$ is meaning, and $z_k$ is linearly convergent to $z_0$, thereby $x_k$ and $\lambda_k$ are linearly convergent to $x$ and $\lambda$, respectively.

**Proof:** From Lemma 3, we only prove that $v$ and $B_k$ satisfy conditions R1, R2, R3 and the inequality (6). From assumption A1, it is obvious that $v$ is continuously differentiable. From A3, it holds that

$$v'(z_k) = v(x_k, \lambda_k) = 0,$$

and

$$v'(z_k) = -\begin{pmatrix} \nabla^2 f(x_k) + \nabla^2 g(x_k) \lambda_k & \nabla g(x_k) \\ \nabla g(x_k)^T & 0 \end{pmatrix}.$$
From assumption A4, it holds that \( v' \) is Lipschitz continuous at \( z_{\ast} \). In a word, \( v \) satisfies the conditions R1, R2, R3.

Now, we prove that, for \( v(z_{\ast}) \) and \( B_{k} \) generated according to (5), (6) is satisfied.

In fact, from (5), we have

\[
B_{k+1} - v'(z_{\ast}) = B_{k} - v'(z_{\ast}) + \left( \frac{q_{k} - B_{k} p_{k}}{p_{k}' p_{k}} \right) p_{k}'
\]

So, \( B_{k+1} - v'(z_{\ast}) \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

Frobenius norm. From (8),

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

So, \( \| D_{k+1} \|_{F} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

Therefore, \( \| D_{k+1} \|_{F} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

Since

\[
\| D_{k} \|_{F} = \left\| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right\|_{F} + \left| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right|_{F}.
\]

So, \( \| D_{k+1} \|_{F} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

The claim holds.

**Theorem 2** Under above-mentioned assumptions, if \( \epsilon, \delta \) in Theorem 1 satisfy that

\[
6\delta \left\| \left( v'(z_{\ast}) \right)' \right\| < 1, 3\epsilon \delta \leq 2\delta,
\]

then \( \left( x_{k} \right) \) is superlinearly convergent to \( \left( x_{\ast} \right) \), i.e.,

\[
\left\| x_{k+1} - x_{\ast} \right\|_{2} = o \left( \left\| x_{k} - x_{\ast} \right\|_{2} \right).
\]

Proof. From Lemma 3, we only prove that \( v \) and \( B_{k} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

Frobenius norm. From (8),

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

Since

\[
\| D_{k} \|_{F} = \left\| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right\|_{F} + \left| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right|_{F}.
\]

So, \( \| D_{k+1} \|_{F} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

The claim holds.

**Theorem 2** Under above-mentioned assumptions, if \( \epsilon, \delta \) in Theorem 1 satisfy that

\[
6\delta \left\| \left( v'(z_{\ast}) \right)' \right\| < 1, 3\epsilon \delta \leq 2\delta,
\]

then \( \left( x_{k} \right) \) is superlinearly convergent to \( \left( x_{\ast} \right) \), i.e.,

\[
\left\| x_{k+1} - x_{\ast} \right\|_{2} = o \left( \left\| x_{k} - x_{\ast} \right\|_{2} \right).
\]

Proof. From Lemma 3, we only prove that \( v \) and \( B_{k} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

Frobenius norm. From (8),

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

Since

\[
\| D_{k} \|_{F} = \left\| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right\|_{F} + \left| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right|_{F}.
\]

So, \( \| D_{k+1} \|_{F} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

The claim holds.

**Theorem 2** Under above-mentioned assumptions, if \( \epsilon, \delta \) in Theorem 1 satisfy that

\[
6\delta \left\| \left( v'(z_{\ast}) \right)' \right\| < 1, 3\epsilon \delta \leq 2\delta,
\]

then \( \left( x_{k} \right) \) is superlinearly convergent to \( \left( x_{\ast} \right) \), i.e.,

\[
\left\| x_{k+1} - x_{\ast} \right\|_{2} = o \left( \left\| x_{k} - x_{\ast} \right\|_{2} \right).
\]

Proof. From Lemma 3, we only prove that \( v \) and \( B_{k} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

Frobenius norm. From (8),

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

Since

\[
\| D_{k} \|_{F} = \left\| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right\|_{F} + \left| \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right|_{F}.
\]

So, \( \| D_{k+1} \|_{F} \) satisfy (7). Denote \( D_{k} = B_{k} - v'(z_{\ast}) \), \( \| D_{k} \|_{F} \) norm.

\[
\| D_{k+1} \|_{F} \leq \left| D_{k} \left( I - \frac{p_{k} p_{k}'}{p_{k}' p_{k}} \right) \right| + \frac{\| q_{k} - v'(z_{\ast}) p_{k} \|}{p_{k}' p_{k}}.
\]

The claim holds.
Lemma 4 Let \( R : R^n \to R^n \) be a function defined by
\[
R(x) = A(x)\left(\nabla f(x) + \nabla g(x) \lambda \right) + \nabla g(x)g(x),
\]
where \( A(x) = I_n - \nabla g(x) \left(\nabla g(x)^\top \nabla g(x)\right)^{-1} \nabla g(x)^\top \).

Then, \( \|x_{k+1} - x_t\| = o(\|x_{k+1} - x_t\|) \) if and only if
\[
\|R(x_{k+1})\| = o(\|x_{k+1} - x_t\|).
\]

Proof. Obviously, using the triangle inequality, it holds that
\[
\|x_{k+1} - x_t\| = o(\|x_{k+1} - x_t\|) \Leftrightarrow \|x_{k+1} - x_t\| = o(\|x_{k+1} - x_t\|).
\]

In addition, from A1, we know that \( R(x) \) is continuously differentiable, and
\[
R(x) = 0,
\]
\[
R'(x) = A(x)\nabla^2_{xx}L(x_t, \lambda_t) + \nabla g(x)\nabla g(x)^\top \quad (15)
\]

It holds that \( R'(x) \) is nonsingular. In fact, let \( d \neq 0 \), and
\[
\left( A(x)\nabla^2_{xx}L(x_t, \lambda_t) + \nabla g(x)\nabla g(x)^\top \right) d = 0,
\]
the facts that \( \nabla g(x) \) has full rank, \( A(x) \) is semi-positive definite and \( \nabla^2_{xx}L(x_t, \lambda_t) \) is positive definite imply that
\[
A(x)\nabla^2_{xx}L(x_t, \lambda_t) d = 0, \nabla g(x)^\top d = 0.
\]
So,
\[
A(x) d = d^\top \nabla^2_{xx}L(x_t, \lambda_t) d = 0,
\]
which is a contradiction.

Thereby, from A4 and Lemma 2, there exist
\[
\beta > \alpha > 0 \quad \text{such that} \quad \alpha \|x_{k+1} - x_t\| \leq \|R(x_{k+1})\| \leq \beta \|x_{k+1} - x_t\|.
\]
So,
\[
\|R(x_{k+1})\| = o(x_{k+1} - x_t) \Leftrightarrow \|x_{k+1} - x_t\| = o(\|x_{k+1} - x_t\|).
\]

The claim holds.

Theorem 3 Under above-mentioned conditions, \( \{x_k\} \) is superlinearly convergent to \( x_t \) if and only if
\[
\left| A(x)\left(\nabla f(x_t) + \nabla g(x_t) \lambda_t \right) + \left(\nabla^2 f(x_t) + \nabla^2 g(x_t) \lambda_t \right)(x_{k+1} - x_t) \right|
\]
\[
= o(\|x_{k+1} - x_t\|),
\]
and
\[
\|g(x_t) + \nabla g(x_t)^\top (x_{k+1} - x_t)\| = o(\|x_{k+1} - x_t\|), \quad \text{i.e.,}
\]
\[
\left| A(x_t) \begin{bmatrix} 0 \ 0 \end{bmatrix} \right| + \begin{bmatrix} v(z_t) \ v'(z_t)(x_{k+1} - x_t) \end{bmatrix}
\]
\[
= o(\|x_{k+1} - x_t\|).
\]

Proof. The sufficiency: Suppose that (16) holds. We only prove that \( \|R(x_{k+1})\| = o(\|x_{k+1} - x_t\|) \).

From (4), we have
\[
v(x_t, \lambda_t) = B_k p_k + \begin{bmatrix} \nabla g(x_t) \lambda_t \ - \lambda_t \end{bmatrix},
\]

So, the fact \( A(x_t) \nabla g(x_t) = 0 \) implies that
\[
(1-A(x_t) 0_{oem})B_k p_k - (A(x_t) - A(x_t) 0_{oem})v(x_t, \lambda_t) = 0,
\]
thus,
\[
-R(x_{k+1}) = (A(x_t) 0_{oem})B_k p_k - (R(x_{k+1}) - R(x_t))
\]
\[
- \nabla g(x_t) g(x_t).
\]

From (15), we have
\[
R'(x_t)(x_{k+1} - x_t)
\]
\[
= \left( A(x_t)\nabla^2_{xx}L(x_t, \lambda_t) + \nabla g(x_t)\nabla g(x_t)^\top \right)(x_{k+1} - x_t)
\]
\[
= -A(x_t) 0_{oem})v(z_t) p_k + \nabla g(x_t)\nabla g(x_t)^\top (x_{k+1} - x_t)
\]
So,
\[
-R(x_{k+1}) = (R(x_{k+1}) - R_k - R'(x_t)(x_{k+1} - x_t))
\]
\[
- \nabla g(x_t) g(x_t) + \nabla g(x_t)^\top (x_{k+1} - x_t)
\]
\[
+ (A(x_t) 0_{oem})v(z_t)
\]
\[
+ (A(x_t) 0_{oem})v'(z_t) p_k.
\]

While
\[
(A(x_t) 0_{oem})v(z_t) + (A(x_t) 0_{oem})v'(z_t) p_k
\]
\[
= -A(x_t)\left(\nabla f(x_t) + \nabla g(x_t) \lambda_t \right)
\]
\[
- A(x_t)\left(\nabla^2 f(x_t) + \nabla^2 g(x_t) \lambda_t \right)(x_{k+1} - x_t)
\]
\[
= -A(x_t)\left(\nabla f(x_t) + \nabla g(x_t) \lambda_t \right)
\]
\[
+ \left(\nabla^2 f(x_t) + \nabla^2 g(x_t) \lambda_t \right)(x_{k+1} - x_t)
\]
\[
+ (A(x_t) - A(x_t))\left(\nabla f(x_t) + \nabla g(x_t) \lambda_t \right)(x_{k+1} - x_t).
\]

In addition, according to Lemma 1, we have
\[
\|R(x_{k+1}) - R_k - R'(x_t)(x_{k+1} - x_t)\| = o(\|x_{k+1} - x_t\|),
\]
so, from (16), it holds that \( \|R(x_{k+1})\| = o(\|x_{k+1} - x_t\|) \).

The necessity: Suppose that \( \|x_{k+1} - x_t\| = o(\|x_{k+1} - x_t\|) \).

On one hand,
\[
g(x_t) + \nabla g(x_t)^\top (x_{k+1} - x_t)
\]
\[
= g(x_t) + \nabla g(x_t)^\top (x_{k+1} - x_t)
\]
\[
+ \left(\nabla g(x_t)^\top - \nabla g(x_t)^\top \right)(x_{k+1} - x_t).
\]
So, in order to prove that 
\[ \| g(x_k) + \nabla g(x_k)^T(x_{k+1} - x_k) \| = o(\|x_{k+1} - x_k\|). \]

While
\[
\| g(x_k) + \nabla g(x_k)^T(x_{k+1} - x_k) \|
= \left( g(x_{k+1}) - g(x_k) - \nabla g(x_k)^T(x_{k+1} - x_k) \right)
+ \left( g(x_k) - g(x_k) \right).
\]
According to Lemma 1, we have
\[ g(x_{k+1}) - g(x_k) - \nabla g(x_k)^T(x_{k+1} - x_k) = o(\|x_{k+1} - x_k\|). \]

On the other hand, the fact implies that
\[ g(x_k) + \nabla g(x_k)^T(x_{k+1} - x_k) = o(\|x_{k+1} - x_k\|). \]

3. Acknowledgements

This work was supported in part by NNSF (No. 1106011) of China and Guangxi Fund for Distinguished Young Scholars (2012GXSFFA060003).

REFERENCES


