

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 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 In this paper, it is proposed to consider the following nonlinear mathematical programming problem:

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

where $f: R^n \rightarrow R, g_j (j \in I): R^n \rightarrow R$ are continuously differentiable functions. Denote the feasible set as follows:

$$X = \{x \in R^n \mid g_j(x) = 0, j \in I\}.$$

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

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

If (x_*, λ_*) is a KKT point pair of Equation (1), then $\nabla Lx, \lambda \Big|_{x=x_*, \lambda=\lambda_*} = 0, i.e.,$

$$\begin{aligned} \nabla f(x_*) + \nabla g(x_*) \lambda_* &= 0, \\ g(x_*) &= 0, \end{aligned} \quad (2)$$

where $g(x_*) = (g(x_*), j \in I), \nabla g(x_*) = (\nabla g(x_*), j \in I)$.

At the point pair (x_k, λ_k) , the Newton's iteration of (2) is defined as follows:

$$\begin{aligned} &\begin{pmatrix} \nabla_{xx}^2 L(x_k, \lambda_k) & \nabla g(x_k) \\ \nabla g(x_k)^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x_k \\ \Delta \lambda_k \end{pmatrix} \\ &= - \begin{pmatrix} \nabla f(x_k) + \nabla g(x_k) \lambda_k \\ g(x_k) \end{pmatrix} \end{aligned} \quad (3)$$

Later, a positive definite matrix H_k^{-1} is replaced for $\nabla_{xx}^2 L(x_k, \lambda_k)$ 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

$$\begin{pmatrix} \nabla_{xx}^2 L(x_k, \lambda_k) & \nabla g(x_k) \\ \nabla g(x_k)^T & 0 \end{pmatrix}.$$

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

$$\begin{aligned} z &= \begin{pmatrix} x \\ \lambda \end{pmatrix}, z_k = \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix}, B_k = \begin{pmatrix} \nabla_{xx}^2 L(x_k, \lambda_k) & \nabla g(x_k) \\ \nabla g(x_k)^T & 0 \end{pmatrix}, \\ v(z_k) &= - \begin{pmatrix} \nabla f(x_k) + \nabla g(x_k) \lambda_k \\ g(x_k) \end{pmatrix}, P_k = z_{k+1} - z_k = \begin{pmatrix} \Delta x_k \\ \Delta \lambda_k \end{pmatrix}, \\ q_k &= v(z_{k+1}) - v(z_k), \\ F(z) &= \|\nabla f(x) + \nabla g(x) \lambda\|^2 + \|g(x)\|^2. \end{aligned}$$

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

$$B_k p_k = v(z_k). \quad (4)$$

To the system of linear Equations (4), like unconstrained optimization, B_k is dealt with using Broyden rank one modifications as follows:

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

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

Algorithm A:

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

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 $x \in R^n$, 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_*, λ_*) , such that $\nabla_{xx}^2 L(x_*, \lambda_*)$ is positive definite;

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

Lemma 1 [9] Let $F: R^{n+m} \rightarrow R^{n+m}$ 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

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

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

$$\|F(u) - F(v) - F'(u-v)\| \leq v \frac{\|u-x\| + \|v-x\|}{2} \|u-v\|.$$

Lemma 2 [9] Let $F: R^{n+m} \rightarrow R^{n+m}$ 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 $F: R^{n+m} \rightarrow R^{n+m}$, which satisfies:

R1 F is continuously differentiable on D ;

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

R3 F' is Lipschitz continuous at z_* , i.e., there exists a constant v , such that

$$\|F'(z_*) - F'(z_*)\| \leq v \|z - z_*\|, z \in D.$$

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

$$\begin{aligned} & \|H_{k+1} - F'(z_*)\| \\ & \leq \|H_k - F'(z_*)\| + \frac{V}{2} (\|z_{k+1} - z_*\| + \|z_k - z_*\|), \forall k, \end{aligned} \quad (6)$$

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

$$\|z_0 - z_*\| < \varepsilon, \|H_0 - F'(z_*)\| < \delta,$$

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

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

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

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

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

then $\{z_k\}$ is meaning, and z_k is linearly convergent to z_* , thereby x_k and λ_k are linearly convergent to x_* and λ_* 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

$$v(z_*) = v(x_*, \lambda_*) = 0, \text{ and}$$

$$v'(z_*) = - \begin{pmatrix} \nabla^2 f(x_*) + \nabla^2 g(x_*) \lambda_* & \nabla g(x_*) \\ \nabla g(x_*)^T & 0 \end{pmatrix}.$$

While $\nabla_{xx}^2 L(x_*, \lambda_*) = \nabla^2 f(x_*) + \nabla^2 g(x_*) \lambda_*$ is positive definite, and $\{\nabla g_j(x_*), j \in I\}$ are linearly independent. So, it is easy to see that $v'(z_*)$ is reversible. In addition,

$$v'(z_k) = - \begin{pmatrix} \nabla^2 f(x_k) + \nabla^2 g(x_k) \lambda_k & \nabla g(x_k) \\ \nabla g(x_k)^T & 0 \end{pmatrix}.$$

From assumption A4, it holds that v' is Lipschitz continuous at z_* . In a word, v satisfies the conditions R1, R2, R3.

Now, we prove that, for $v(z_*)$ and B_k generated according to (5), (6) is satisfied.

In fact, from (5), we have

$$\begin{aligned} & B_{k+1} - v'(z_*) \\ &= B_k - v'(z_*) + \frac{(q_k - B_k p_k) p_k^T}{p_k^T p_k} \\ &= B_k - v'(z_*) + \frac{(v'(z_*) p_k - B_k p_k) p_k^T}{p_k^T p_k} \\ &\quad + \frac{(q_k - v'(z_*) p_k) p_k^T}{p_k^T p_k} \\ &= B_k - v'(z_*) \left(I - \frac{p_k p_k^T}{p_k^T p_k} \right) + \frac{(q_k - v'(z_*) p_k) p_k^T}{p_k^T p_k}. \end{aligned} \tag{8}$$

So,

$$\begin{aligned} & \|B_{k+1} - v'(z_*)\| \\ &\leq \|B_k - v'(z_*)\| \left\| I - \frac{p_k p_k^T}{p_k^T p_k} \right\| + \frac{\|q_k - v'(z_*) p_k\|}{\|p_k\|}. \end{aligned}$$

While

$$\left\| I - \frac{p_k p_k^T}{p_k^T p_k} \right\| = \frac{p_k^T p_k}{p_k^T p_k} = 1,$$

and from Lemma 1, we have

$$\begin{aligned} & \|q_k - v'(z_*) p_k\| \\ &= \|v(z_{k+1}) - v(z_k) - v'(z_*) (z_{k+1} - z_k)\| \\ &\leq \frac{\nu}{2} (\|z_{k+1} - z_*\| + \|z_k - z_*\|) \|z_{k+1} - z_k\|. \end{aligned} \tag{9}$$

So,

$$\|B_{k+1} - v'(z_*)\| \leq \|B_k - v'(z_*)\| + \frac{\nu}{2} (\|z_{k+1} - z_*\| + \|z_k - z_*\|),$$

i.e., (6) is true, thereby, from Lemma 3, we get

$$z_k \rightarrow z_*, \text{ i.e., } x_k \rightarrow x_*, \lambda_k \rightarrow \lambda_*.$$

The claim holds.

Theorem 2 Under above-mentioned assumptions, if ε, δ in Theorem 1 satisfy that

$$6\delta \|(v'(z_*))^{-1}\| < 1, 3\nu\varepsilon \leq 2\delta, \tag{10}$$

then $\left\{ \begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} \right\}$ is superlinearly convergent to $\begin{pmatrix} x_* \\ \lambda_* \end{pmatrix}$, i.e.,

$$\left\| \begin{pmatrix} x_{k+1} - x_* \\ \lambda_{k+1} - \lambda_* \end{pmatrix} \right\| = o\left(\left\| \begin{pmatrix} x_k - x_* \\ \lambda_k - \lambda_* \end{pmatrix} \right\| \right).$$

Proof. From Lemma 3, we only prove that v and B_k

satisfy (7). Denote $D_k = B_k - v'(z_*)$, $\|\cdot\|_2$ norm, $\|\cdot\|_F$ Frobenius norm. From (8),

$$\|D_{k+1}\|_F \leq \left\| D_k \left(I - \frac{p_k p_k^T}{p_k^T p_k} \right) \right\|_F + \frac{\|(q_k - v'(z_*) p_k) p_k^T\|_F}{p_k^T p_k}. \tag{11}$$

Since

$$\begin{aligned} \left\| D_k \frac{p_k p_k^T}{p_k^T p_k} \right\|_F^2 &= \text{tr} \left(\frac{p_k p_k^T D_k^T p_k p_k^T}{(p_k^T p_k)^2} \right) \\ &= \frac{(D_k p_k)^T (D_k p_k) p_k^T p_k}{(p_k^T p_k)^2} = \frac{\|D_k p_k\|_F^2}{\|p_k\|_F^2}, \end{aligned}$$

$$\begin{aligned} \|D_k\|_F^2 &= \text{tr}(D_k^T D_k) = \left\| D_k \frac{p_k p_k^T}{p_k^T p_k} \right\|_F^2 + \left\| D_k \left(I - \frac{p_k p_k^T}{p_k^T p_k} \right) \right\|_F^2 \\ &= \frac{\|D_k p_k\|_F^2}{\|p_k\|_F^2} + \left\| D_k \left(I - \frac{p_k p_k^T}{p_k^T p_k} \right) \right\|_F^2. \end{aligned}$$

So,

$$\begin{aligned} \left\| D_k \left(I - \frac{p_k p_k^T}{p_k^T p_k} \right) \right\|_F &= \left(\|D_k\|_F^2 - \frac{\|D_k p_k\|_F^2}{\|p_k\|_F^2} \right)^{\frac{1}{2}} \\ &\leq \|D_k\|_F - \frac{1}{2\|D_k\|_F} \left(\frac{\|D_k p_k\|_F}{\|p_k\|_F} \right)^2. \end{aligned} \tag{12}$$

In addition, from (6), (10), using the method of mathematical induction, it is not difficult to prove that

$$\|B_k - v'(z_*)\| \leq (2 - 2^{-k})\delta, \|z_{k+1} - z_*\| \leq \frac{1}{2} \|z_k - z_*\|, \forall k, \tag{13}$$

thereby,

$$\|D_k\|_F \leq 2\delta, \sum_{k=0}^{\infty} \|z_k - z_*\| \leq 2\varepsilon, \forall k. \tag{14}$$

Thus, from (9), (12), (13), we have

$$\begin{aligned} \frac{\|D_k p_k\|_F^2}{\|p_k\|_F^2} &\leq 2\|D_k\|_F \left(\|D_k\|_F - \|D_{k+1}\|_F + \frac{3}{4}\nu \|z_k - z_*\| \right), \\ \sum_{k=0}^i \frac{\|D_k p_k\|_F^2}{\|p_k\|_F^2} &\leq 4\delta \left(\|D_0\|_F - \|D_{i+1}\|_F + \frac{3}{4}\nu \sum_{k=0}^i \|z_k - z_*\| \right) \\ &\leq 4\delta^2 + 6\nu\varepsilon, \forall i, \end{aligned}$$

i.e. $\sum_{k=0}^i \frac{\|D_k p_k\|_F^2}{\|p_k\|_F^2}$ is convergent, thereby, $\lim_{k \rightarrow \infty} \frac{\|D_k p_k\|_F^2}{\|p_k\|_F^2} = 0$.

From Theorem 2, we don't conclude that $\{x_k\}$ is superlinearly convergent to x_* , i.e.,

$\|x_{k+1} - x_*\| = o(\|x_k - x_*\|)$. In the sequel, we discuss one condition which assures that x_k is superlinearly conver-

gent to x_* .

Lemma 4 Let $R: R^n \rightarrow R^n$ be a function defined by

$$R(x) = A(x)(\nabla f(x) + \nabla g(x)\lambda_*) + \nabla g(x_*)g(x),$$

where $A(x) = I_n - \nabla g(x)(\nabla g(x)^{-1}\nabla g(x))^{-1}\nabla g(x)^T$.

Then, $\|x_{k+1} - x_*\| = o(\|x_k - x_*\|)$, if and only if

$$\|R(x_{k+1})\| = o(\|x_{k+1} - x_k\|).$$

Proof. Obviously, using the triangle inequality, it holds that

$$\|x_{k+1} - x_*\| = o(\|x_k - x_*\|) \Leftrightarrow \|x_{k+1} - x_*\| = o(\|x_{k+1} - x_k\|).$$

In addition, from A1, we know that $R(x)$ is continuously differentiable, and

$$\begin{aligned} R(x_*) &= 0, \\ R'(x_*) &= A(x_*)\nabla_{xx}^2 L(x_*, \lambda_*) + \nabla g(x_*)\nabla g(x_*)^T. \end{aligned} \quad (15)$$

It holds that $R'(x_*)$ is nonsingular. In fact, let $d \neq 0$, and

$$\left(A(x_*)\nabla_{xx}^2 L(x_*, \lambda_*) + \nabla g(x_*)\nabla g(x_*)^T \right) d = 0,$$

the facts that $\nabla g(x_*)$ has full rank, $A(x_*)$ is semi-positive definite and $\nabla_{xx}^2 L(x_*, \lambda_*)$ is positive definite imply that

$$A(x_*)\nabla_{xx}^2 L(x_*, \lambda_*)d = 0, \nabla g(x_*)^T d = 0.$$

So,

$$A(x_*)d = d, d^T \nabla_{xx}^2 L(x_*, \lambda_*)d = 0,$$

which is a contradiction.

Thereby, from A4 and Lemma 2, there exist $\beta > \alpha > 0$, such that

$$\alpha \|x_{k+1} - x_*\| \leq \|R(x_{k+1})\| \leq \beta \|x_{k+1} - x_*\|.$$

So,

$$\|R(x_{k+1})\| = o(\|x_{k+1} - x_k\|) \Leftrightarrow \|x_{k+1} - x_*\| = o(\|x_{k+1} - x_k\|).$$

The claim holds.

Theorem 3 Under above-mentioned conditions, $\{x_k\}$ is superlinearly convergent to x_* if and only if

$$\begin{aligned} & \left\| A(x_k)(\nabla f(x_k) + \nabla g(x_k)\lambda_k) \right. \\ & \left. + (\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k) \right\| \\ & = o(\|x_{k+1} - x_k\|), \end{aligned}$$

and $\left\| g(x_k) + \nabla g(x_*)^T(x_{k+1} - x_k) \right\| = o(\|x_{k+1} - x_k\|)$, i.e.,

$$\begin{aligned} & \left\| \begin{pmatrix} A(x_k) & 0_{n \times m} \\ 0_{m \times n} & I_{m \times m} \end{pmatrix} (v(z_k) + v'_x(z_*)(x_{k+1} - x_k)) \right\| \\ & = o(\|x_{k+1} - x_k\|). \end{aligned} \quad (16)$$

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

From (4), we have

$$v(x_k, \lambda_*) = B_k p_k + \begin{pmatrix} \nabla g(x_k)(\lambda_k - \lambda_*) \\ 0 \end{pmatrix},$$

So, the fact $A(x_k)\nabla g(x_k) = 0$ implies that

$$\left(A(x_k) \quad 0_{n \times m} \right) B_k p_k - \left(A(x_k) \quad 0_{n \times m} \right) v(x_k, \lambda_*) = 0,$$

thereby,

$$\begin{aligned} -R(x_{k+1}) &= \left(A(x_k) \quad 0_{n \times m} \right) B_k p_k - \left(R(x_{k+1}) - R(x_k) \right) \\ & \quad - \nabla g(x_*)g(x_k). \end{aligned}$$

From (15), we have

$$\begin{aligned} & R'(x_*)(x_{k+1} - x_k) \\ &= \left(A(x_*)\nabla_{xx}^2 L(x_*, \lambda_*) + \nabla g(x_*)\nabla g(x_*)^T \right) (x_{k+1} - x_k) \\ &= -\left(A(x_k) \quad 0_{n \times m} \right) v'(z_*)p_k + \nabla g(x_*)\nabla g(x_*)^T (x_{k+1} - x_k) \\ & \text{So,} \end{aligned}$$

$$\begin{aligned} -R(x_{k+1}) &= \left(R(x_{k+1}) - R_k - R'(x_*)(x_{k+1} - x_k) \right) \\ & \quad - \nabla g(x_*) \left(g(x_k) + \nabla g(x_*)^T (x_{k+1} - x_k) \right) \\ & \quad + \left(A(x_k) \quad 0_{n \times m} \right) v(z_k) \\ & \quad + \left(A(x_k) \quad 0_{n \times m} \right) v'(z_k)p_k. \end{aligned}$$

While

$$\begin{aligned} & \left(A(x_k) \quad 0_{n \times m} \right) v(z_k) + \left(A(x_k) \quad 0_{n \times m} \right) v'(z_k)p_k \\ &= -A(x_k)(\nabla f(x_k) + \nabla g(x_k)\lambda_k) \\ & \quad - A(x_*)(\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k) \\ &= -A(x_k)(\nabla f(x_k) + \nabla g(x_k)\lambda_k) \\ & \quad + (\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k) \\ & \quad + (A(x_k) - A(x_*))(\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k). \end{aligned}$$

In addition, according to Lemma 1, we have

$$\|R(x_{k+1}) - R_k - R'(x_*)(x_{k+1} - x_k)\| = o(\|x_{k+1} - x_k\|),$$

so, from (16), it holds that $\|R(x_{k+1})\| = o(\|x_{k+1} - x_k\|)$.

The necessity: Suppose that $\|x_{k+1} - x_*\| = o(\|x_k - x_*\|)$ i.e., $\|x_{k+1} - x_*\| = o(\|x_{k+1} - x_k\|)$.

On one hand,

$$\begin{aligned} & g(x_k) + \nabla g(x_*)^T(x_{k+1} - x_k) \\ &= g(x_k) + \nabla g(x_k)^T(x_{k+1} - x_k) \\ & \quad + (\nabla g(x_*)^T - \nabla g(x_k)^T)(x_{k+1} - x_k). \end{aligned}$$

So, in order to prove that

$$\left\|g(x_k) + \nabla g(x_k)^T(x_{k+1} - x_k)\right\| = o(\|x_{k+1} - x_k\|).$$

While

$$\begin{aligned} & \left\|g(x_k) + \nabla g(x_k)^T(x_{k+1} - x_k)\right\| \\ &= -\left(g(x_{k+1}) - g(x_k) - \nabla g(x_k)^T(x_{k+1} - x_k)\right) \\ & \quad + \left(g(x_{k+1}) - g(x_*)\right). \end{aligned}$$

According to Lemma 1, we have

$$\begin{aligned} & g(x_{k+1}) - g(x_k) - \nabla g(x_k)^T(x_{k+1} - x_k) \\ &= o(\|x_{k+1} - x_k\|), \end{aligned}$$

and

$$\begin{aligned} & \left(g(x_{k+1}) - g(x_k)\right) \\ & \leq k_1 \|x_{k+1} - x_*\| = o(\|x_{k+1} - x_k\|) \quad (k_1 > 0). \end{aligned}$$

So, $\left\|g(x_k) + \nabla g(x_k)^T(x_{k+1} - x_k)\right\| = o(\|x_{k+1} - x_k\|).$

On the other hand, the fact $A(x_k)\nabla g(x_k) = 0$ implies that

$$\begin{aligned} & A(x_k)(\nabla f(x_k) + \nabla g(x_k)\lambda_k) \\ & \quad + (\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k) \\ &= A(x_k)(\nabla f(x_k) + \nabla g(x_k)\lambda_*) \\ & \quad + (\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k). \end{aligned}$$

Let $h_{\lambda_*}(x) = \nabla f(x) + \nabla g(x)\lambda_*$, then

$$\begin{aligned} & \nabla f(x_k) + \nabla g(x_k)\lambda_k + (\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k) \\ &= h_{\lambda_*}(x_k) + \nabla h_{\lambda_*}(x_*)^T(x_{k+1} - x_k) \\ &= -\left(h_{\lambda_*}(x_{k+1}) - h_{\lambda_*}(x_k) - \nabla h_{\lambda_*}(x_*)^T(x_{k+1} - x_k)\right) \\ & \quad + \left(h_{\lambda_*}(x_{k+1}) - h_{\lambda_*}(x_*)\right). \end{aligned}$$

It is easy to see that

$$\left\|h_{\lambda_*}(x_k) + \nabla h_{\lambda_*}(x_*)^T(x_{k+1} - x_k)\right\| = o(\|x_{k+1} - x_k\|).$$

Thereby,

$$\begin{aligned} & \left\|A(x_k)(\nabla f(x_k) + \nabla g(x_k)\lambda_k)(\nabla f(x_k) + \nabla g(x_k)\lambda_k) \right. \\ & \quad \left. + (\nabla^2 f(x_*) + \nabla^2 g(x_*)\lambda_*)(x_{k+1} - x_k)\right\| = o(\|x_{k+1} - x_k\|). \end{aligned}$$

The claim holds.

3. Acknowledgements

This work was supported in part by NNSF (No. 11061011) of China and Guangxi Fund for Distinguished Young Scholars (2012GXSSFFA060003).

REFERENCES

- [1] M. J. D. Powell, "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations," In: G. A. Watson, Ed., *Numerical Analysis*, Springer-Verlag, Berlin, 1978, pp. 144-157. [doi:10.1007/BFb0067703](https://doi.org/10.1007/BFb0067703)
- [2] S.-P. Han, "Superlinearly Convergent Variable Metric Algorithm for General Nolinear Programming Problem," *Mathematical Programming*, Vol. 11, No. 1, 1976, pp. 263-282. [doi:10.1007/BF01580395](https://doi.org/10.1007/BF01580395)
- [3] D. Q. Mayne and E. Polak, "A Superlinearly Convergent Algorithm for Constrained Optimization Problems," *Mathematical Programming Studies*, Vol. 16, 1982, pp. 45-61. [doi:10.1007/BFb0120947](https://doi.org/10.1007/BFb0120947)
- [4] J.-B. Jian, C.-M. Tang, Q.-J. Hu and H.-Y. Zheng, "A Feasible Descent SQP Algorithm for General Constrained Optimization without Strict Complementarity," *Journal of Computational and Applied Mathematics*, Vol. 180, No. 2, 2005, pp. 391-412. [doi:10.1016/j.cam.2004.11.008](https://doi.org/10.1016/j.cam.2004.11.008)
- [5] Z. Jin and Y. Q. Wang, "A Simple Feasible SQP Method for Inequality Constrained Optimization with Global and Superlinear Convergence," *Journal of Computational and Applied Mathematics*, Vol. 233, No. 1, 2010, pp. 3060-3073. [doi:10.1016/j.cam.2009.11.061](https://doi.org/10.1016/j.cam.2009.11.061)
- [6] Y.-F. Yang, D.-H. Li and L. Q. Qi, "A Feasible Sequential Linear Equation Method for Inequality Constrained Optimization," *SIAM Journal on Optimization*, Vol. 13, No. 4, 2003, pp. 1222-1244. [doi:10.1137/S1052623401383881](https://doi.org/10.1137/S1052623401383881)
- [7] Z. B. Zhu, "An Interior Point Type QP-Free Algorithm with Superlinear Convergence for Inequality Constrained Optimization," *Applied Mathematical Modelling*, Vol. 31, No. 6, 2007, pp. 1201-1212. [doi:10.1016/j.apm.2006.04.019](https://doi.org/10.1016/j.apm.2006.04.019)
- [8] J. B. Jian, D. L. Han and Q. J. Xu, "A New Sequential Systems of Linear Equations Algorithm of Feasible Descent for Inequality Constrained Optimization," *Acta Mathematica Sinica, English Series*, Vol. 26, No. 12, 2010, pp. 2399-2420. [doi:10.1007/s10114-010-7432-0](https://doi.org/10.1007/s10114-010-7432-0)
- [9] Y. Yuan and W. Sun, "Theory and Methods of Optimization," Science Press, Beijing, 1997.