A NOTE ON THE COMPARISON OF THE KANTOROVICH AND MOORE THEOREMS

SHEN ZUHE*
Department of Mathematics, Nanjing University, People's Republic of China

and

M. A. WOLFE
Department of Mathematical Sciences, University of St Andrews, Fife KY16 9SS, Scotland

(Received 6 April 1989; received for publication 16 November 1989)

Key words and phrases: Affine invariant form of the Kantorovich theorem, Moore's test, interval analysis.

1. INTRODUCTION

Consider the finite system of nonlinear algebraic equations

\[ F(x) = 0 \]  

(1.1)

where \( F: D \subseteq \mathbb{R}^n \to \mathbb{R}^n \) is continuously differentiable on the open convex set \( D \). A theoretical comparison by Rall [6] of the theorems of Kantorovich [2] and of Moore [3, 4] which contain sufficient conditions for the existence of a unique zero of \( F \) in a given subset of \( D \) shows that the Kantorovich theorem has only a slight advantage over the Moore theorem with regard to sensitivity and precision, while the latter requires far less computational labour than the former.

Rall's comparison is based on the assumption that the interval extension \( F': I(D) \to I(\mathbb{R}^n) \) of the derivative \( F: D \to \mathbb{R}^n \) is defined by

\[ F'(z) = [F'(x) \mid x \in z]. \]

Recently Neumaier and Shen [5] have shown that when the derivative in the Krawczyk operator is replaced with a suitable slope then the corresponding existence theorem is always at least as effective as the Kantorovich theorem with respect to sensitivity and precision. Deuflhard and Heindl [1] have given an affine invariant form of the Kantorovich theorem containing sufficient conditions for existence and uniqueness which are weaker than those of the original Kantorovich theorem. It is shown in this paper that the Moore theorem is easily made affine invariant and that the affine invariant form of the Moore theorem compares favourably with the affine invariant form of the Kantorovich theorem due to Deuflhard and Heindl.

2. AN AFFINE INVARIANT FORM OF THE KANTOROVICH THEOREM

As pointed out by Deuflhard and Heindl [1], Newton's method for the solution of (1.1) is affine invariant in that if \( A \in \mathbb{R}^{n \times n} \) is nonsingular, \( G: D \subseteq \mathbb{R}^n \to \mathbb{R}^n \) is define by

\[ G(x) = AF(x) \]  

(2.1)

*Supported by SERC visiting fellowship: GR/F 11092.
and the sequences $(x^{(k)})$ and $(y^{(k)})$ are generated from

$$x^{(k+1)} = x^{(k)} - F'(x^{(k)})^{-1}F(x^{(k)})$$

(2.2)

and

$$y^{(k+1)} = y^{(k)} - G'(y^{(k)})^{-1}G(y^{(k)})$$

with $y^{(0)} = x(0)$, then $y^{(k)} = x^{(k)}$ ($\forall k \geq 0$), and therefore the theoretical analysis of Newton's method should be affine invariant. Deuflhard and Heindl have proved the following affine invariant form of the Kantorovich theorem.

**Theorem 1.** Let $F: D \subseteq X \to Y$ (where $X, Y$ are real Banach spaces) be Fréchet differentiable on the open convex set $D$. Suppose that $x^{(0)} \in D$ is such that $F'(x^{(0)})^{-1}$ exists and

$$\|F'(x^{(0)})^{-1}F(x^{(0)})\| \leq \alpha,$$  

(2.3)

$$\|F'(x^{(0)})^{-1}[F'(y) - F'(x)]\| \leq \omega \|y - x\| \quad (\forall x, y \in D),$$  

(2.4)

$$h = \alpha \omega \leq \frac{1}{2},$$

where

$$B[x^{(0)}, \rho_-] = \{x \in X \mid \|x - x^{(0)}\| \leq \rho_-\} \subseteq D$$

Then $F'(x)^{-1}$ exists $\forall x \in B(x^{(0)}, \rho_-)$ where

$$B(x^{(0)}, \rho_-) = \{x \in X \mid \|x - x^{(0)}\| < \rho_-\},$$

the sequence $(x^{(k)})$ generated from (2.2) remains in $B(x^{(0)}, \rho_-)$ and $x^{(k)} \to x^*$ ($k \to \infty$) where $F(x^*) = 0$, the zero $x^*$ of $F$ is unique in $B[x^{(0)}, \rho_-] \cup (D \cap B(x^{(0)}, \rho_+))$ where

$$\rho_+ = [1 + \sqrt{1 - 2h}] / \omega,$$

$$\|x^{(k)} - x^*\| \leq \frac{2\sqrt{1 - 2h}}{h} \frac{\theta^2}{1 - \theta^2} \|x^{(1)} - x^{(0)}\| \quad \text{if } h < \frac{1}{2}$$

and

$$\|x^{(k)} - x^*\| \leq 2^{-k+1}\|x^{(1)} - x^{(0)}\| \quad \text{if } h = \frac{1}{2},$$

where $\theta = \rho_- / \rho_+$.

3. **AN AFFINE INVARIANT FORM OF THE MOORE THEOREM**

Suppose that the hypotheses of theorem 1 are valid with $X = Y = R^n$, $x^{(0)} = z = \text{mid}(x_\gamma)$ where $x_\gamma \in I(D)$ is given by

$$x_\gamma = B_\infty[z, \gamma]$$

$$= [z - \gamma e, z + \gamma e]$$

(3.1)

in which $e = (1, \ldots, 1)^T \in R^n$, $\gamma \in R$ and $\gamma > 0$, and let

$$K(x_\gamma) = w - \gamma [F'(z)^{-1}F[z, x_\gamma] - I][-e, e]$$

(3.2)

where

$$w = z - F'(z)^{-1}F(z)$$

(3.3)
and

\[ F[z, \chi] = \int_{0}^{1} F'(z + t(x, - z)) \, dt \]  \hspace{1cm} (3.4)

in which integration is defined as in [4]. Then \( K \) is affine invariant in the sense of (2.1), and Moore’s existence theorem is valid and is affine invariant when \( K \) is defined by (3.3), (3.4). In order to compare the affine invariant Moore theorem with theorem 1 we need the following lemma.

**Lemma 1.**

\[ \|F'(z)^{-1}[F[z, \chi] - F'(z)]\|_\infty \leq \frac{1}{2} \omega \gamma. \]

**Proof.** By (2.4), \( \forall t \in [0, 1] \), \( \forall x \in \chi \),

\[ \|F'(z)^{-1}[F'(z + t(x - z)) - F'(z)]\|_\infty \leq \omega \|x - z\|_\infty, \]

whence

\[
\|F'(z)^{-1}[F[z, \chi] - F'(z)]\|_\infty \leq \int_{0}^{1} \|F'(z)^{-1}[F'(z + t(x, - z)) - F'(z)]\|_\infty \, dt
\]

\[
\leq \int_{0}^{1} \omega \|\chi - z\|_\infty \, dt
\]

\[
= \frac{1}{2} \gamma \omega. \]

The following theorem shows that the affine invariant form of Moore’s theorem is at least as effective as theorem 1 in the sense of Rall [6].

**Theorem 2.** If the hypotheses of theorem 1 are valid with \( x^{(0)} = z = \text{mid}(\chi) \) where \( \chi \in I(D) \) is defined by (3.1) and \( K(\chi) \) is defined by (3.2)-(3.4), then \( K(\chi) \subseteq \chi \).

**Proof.** By (3.3), (3.4) if \( u \in K(\chi) \) then \( u = w + v \) where

\[ v = \gamma[F'(z)^{-1}F[z, \chi] - I][-e, e]. \]

So by lemma 1 and (2.3),

\[ \|z - w\|_\infty \leq \|z - w\|_\infty + \|v\|_\infty \]

\[ \leq \alpha + \gamma \|F'(z)^{-1}[F[z, \chi] - F'(z)]\|_\infty \]

\[ \leq \alpha + \frac{1}{2} \omega \gamma^2. \]

So \( K(\chi) \subseteq \chi \), if

\[
\alpha + \frac{1}{2} \omega \gamma^2 \leq \gamma. \]  \hspace{1cm} (3.5)

Now (3.5) holds if and only if \( \rho_- \leq \gamma \leq \rho_+ \), so if the hypotheses of theorem 1 are valid with \( x^{(0)} = z = \text{mid}(\chi) \) then \( K(\chi) \subseteq \chi \).
4. AN EXAMPLE

Theorem 2 shows that the hypotheses of the affine invariant Moore theorem are always valid when those of theorem 1 are valid. An example which is given in [5] illustrates the fact that the hypotheses of the affine invariant form of the Moore theorem may be satisfied even when those of theorem 1 are not. Let $F: x_1, x_2 \subseteq R^1 \rightarrow R^1$ be defined by

$$F(x) = x^3 - a, \quad a \in [0, 1], \quad x \in x_1 = [a, 2 - a].$$

If $z = \text{mid}(x_1)$ and $h$ is as in theorem 1 then ($\forall a \in [0, \frac{1}{2}]$)

$$h = \frac{3}{4}(1 - a)(2 - a) > \frac{1}{2}.$$ 

So theorem 1 cannot be used to establish the existence of a zero of $F$ in $x_1$ to which the sequence generated from (2.2) converges. On the other hand,

$$K(x_1) = \frac{1}{2}[a^3 - 6a^2 + 10a + 2, -a^3 + 3a^2 - 8a + 6].$$ 

whence if $a \in [0.44, \frac{1}{2})$ then $K(x_1) \subseteq x_1$.

REFERENCES