Reprinted from JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS
All Rights Reserved by Academic Press, New York and London

Vol. 149, No. 2, July 1, 1990
Printed in Belgium

The Krawczyk Operator and Kantorovich's Theorem

ARNOLD NEUMAIER AND SHEN ZUHE*

Institut für Angewandte Mathematik, University Freiburg, Hermann-Herder-Str. 10, D-7800 Freiburg i. Br., West Germany

Submitted by E. Stanley Lee

Received January 17, 1989

R. E. Moore showed that the Krawczyk operator can be used to verify the existence of a unique solution of $F(x^*)=0$ in a hypercube. When the derivative in the Krawczyk operator is replaced by a slope, a comparison shows that the resulting improved form of Moore's existence test is at least as good as and, as shown by an example, sometimes better than that of Kantorovich's theorem with regard to sensitivity and precision.

1. INTRODUCTION

Consider the finite system of nonlinear equation

$$F(x)=0, \tag{1.1}$$

where $F: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be a continuously differentiable function from a domain $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$.

The well-known result of Kantorovich [2] in 1948 and the more recent interval-analytic theorem of Moore [4, 5] based on the Krawczyk operator are both used computationally to establish the existence of a solution x^* of (1.1) and to obtain bounds for the error vector $y-x^*$ of an approximation $y \approx x^*$. A theoretical comparison by Rall [9] between these two theorems on the basis of sensitivity, precision, and computational complexity shows that the Kantorovich theorem has at best only a slight edge in sensitivity and precision over Moore's theorem, while the latter requires far less computational work to apply. Rall's comparison is based on the assumption that the interval extension $F'(x)$ is defined as the range of F' on x. Although it is difficult to compute, it is the theoretically tightest

* This paper has been written while the second author worked as a visiting professor at the Institut für Angewandte Mathematik of the University of Freiburg i. Br./West Germany. It has been sponsored by Stiftung Volkswagenwerk, Number I/63 064. He thanks Professor Dr. K. Nickel who helped him and made his stay possible.

0022-247X/90 $3.00

interval extension of F' in Moore's theorem. Therefore, it has been shown in [9] by an example that this cannot be further improved in general. The purpose of this paper is to show that when the derivative in the Krawczyk operator is replaced by a suitable slope then the corresponding existence test will always be at least as good as that of the Kantorovich theorem in both aspects and often better. Notation and terminology of this paper follow Krawczyk and Neumaier [3] and Neumaier [7]. For convenience of the reader, the theorem by Kantorovich will be restated without proof. We assume in the following:

HYPOTHESIS (K). (i) *Let z be a vector at which the Jacobian matrix $F'(z)$ is invertible with*

$$\|F'(z)^{-1}\| \leqslant \beta.$$

(ii) *Let there be a Lipschitz constant χ for F' such that*

$$\|F'(u)-F'(v)\| \leqslant \chi \|u-v\|$$

for all u, v in a given neighbourhood Ω of z. By (i), *the Newton point*

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

is uniquely defined. Let us put

$$\eta := \|p(z)-z\|. \tag{1.2}$$

1.1. THEOREM (Kantorovich). *Under Hypothesis* (K), *if*

$$h := \beta\chi\eta \leqslant \tfrac{1}{2} \tag{1.3}$$

and $\Omega_0 := \{x \in \mathbb{R}^n \mid \|x-z\| \leqslant \gamma\} \subseteq \Omega$, where

$$\gamma = \frac{1-\sqrt{1-2h}}{h}\eta = \frac{2\eta}{1+\sqrt{1-2h}} \tag{1.4}$$

then (1.1) *has a solution $x^* \in \Omega_0$. Moreover, if $h < \frac{1}{2}$, the solution x^* is unique in Ω_0.*

2. KRAWCZYK OPERATOR AND MOORE TEST

Let $\mathbb{IR}$, $\mathbb{IR}^n$, $\mathbb{IR}^{n\times n}$ denote the set of real intervals, n-dimensional interval vectors, and $n \times n$ interval matrices, and let $\mathbb{I}D := \{x \in \mathbb{IR}^n \mid x \subseteq D\}$ for $D \subseteq \mathbb{R}^n$. Let $x \subseteq \mathbb{IR}^n$. The symbol $\sim$ is often used to denote a generic element $\tilde{x} \in x$. $\check{x} = \frac{1}{2}(\bar{x}+\underline{x})$ denotes the midpoint of an interval (vector) x. The interval hull of a bounded subset $\Sigma \subseteq \mathbb{R}^{n\times n}$ is defined as $\mathbb{I}\Sigma := [\inf(\Sigma), \sup(\Sigma)]$. For properties of interval arithmetic see Moore [6].

The mean value theorem states that, for all $\tilde{x}, z \in x$,

$$F(\tilde{x}) - F(z) = \int_0^1 F'(z+t(\tilde{x}-z))(\tilde{x}-z)\,dt.$$

Hence, if we define

$$F[z,x] := \int_0^1 F'(z+t(x-z))\,dt, \tag{2.1}$$

where $F'(x) = \mathbb{I}\{F'(\tilde{x}) \mid \tilde{x} \in x\}$, then for all $z \in x$,

$$F(\tilde{x}) \in F(z) + F[z,x](x-z), \qquad \tilde{x} \in x. \tag{2.2}$$

Clearly.

$$F[z,x] \subseteq F'(x). \tag{2.3}$$

Therefore, the centered form extension defined as

$$F_s(x) := F(\check{x}) + F[\check{x},x](x-\check{x}) \tag{2.4}$$

yields an enclosure of the range of F on x which is always at least as sharp as that of the mean value form extension (cf. [5]).

Thus, Moore's existence test can be stated in the following improved form, cf. Qi [8], Neumaier [7].

2.1. THEOREM. *If there is a matrix $C \in \mathbb{R}^{n\times n}$ such that the Krawczyk operator*

$$K(x) := z - CF(z) - (CF[z,x]-I)(x-z) \tag{2.5}$$

satisfies $K(x) \subseteq x$, then x contains a zero of (1.1). *Moreover, if $K(x) \subseteq \operatorname{int}(x)$, then x contains a unique zero of* (1.1).

3. COMPARISON BETWEEN THE THEOREMS OF KANTOROVICH AND MOORE

Let $\|\cdot\| = \|\cdot\|_\infty$. For comparison, we suppose that F' satisfies the following matrix Lipschitz condition for some Lipschitz matrix $\Lambda \in \mathbb{R}^{n\times n}$:

$$|F'(\tilde{x}) - F'(\tilde{y})| \leqslant \Lambda |\tilde{x}-\tilde{y}| \qquad \text{for all} \quad \tilde{x}, \tilde{y} \in x. \tag{3.1}$$

Clearly, this implies (ii) of Hypothesis (K) with

$$\chi = \|A\|. \tag{3.2}$$

As in [9], the closed ball Ω_0 in $\mathbb{R}^n_\infty$ with center $z \in \mathbb{R}^n$ and radius γ is the interval

$$\Omega_0 = x_\gamma := [z - \gamma e, z + \gamma e],$$

where $e = (1, 1, ..., 1)^T$. Suppose that $x = x_\gamma$. In this case $z = \check{x}_\gamma$; thus the Krawczyk operator of x_γ is

$$K(x_\gamma) := w - \gamma(CF[z, x_\gamma] - I)[-e, e], \tag{3.3}$$

where

$$w = z - CF(z). \tag{3.4}$$

3.1 LEMMA. *For $x = x_\gamma$, the slope* (2.1) *satisfies*

$$|F[z, x_\gamma] - F'(z)| \leqslant \frac{\gamma}{2} A.$$

Proof. The Lipschitz condition (3.1) implies

$$|F'(z + t(x - z)) - F'(z)| \leqslant tA \|x - z\| \quad \text{for} \quad 0 \leqslant t \leqslant 1.$$

Hence

$$|F[z, x_\gamma] - F'(z)| \leqslant \int_0^1 |F'(z + t(x_\gamma - z)) - F'(z)| \, dt$$
$$\leqslant \int_0^1 t \, dt \cdot A \|x_\gamma - z\|$$
$$= \frac{\gamma}{2} A. \quad \blacksquare$$

3.2 THEOREM. *Let β be a constant with $\beta \geqslant \|C\|$, and let $\eta := \|z - w\|$. If*

$$h = \beta\eta \|A\| \leqslant \tfrac{1}{2}, \tag{3.5}$$

then $K(x_\gamma) \subseteq x_\gamma \subseteq x$ for any γ such that

$$\frac{2\eta}{1 + \sqrt{1 - 2h}} \leqslant \gamma \leqslant \frac{2\eta}{1 - \sqrt{1 - 2h}}. \tag{3.6}$$

Proof. Any element of $K(x_\gamma)$ has the form $w + v$ with

$$v \in \gamma(CF[z, x_\gamma] - I)[-e, e].$$

Since, by Lemma 3.1,

$$|F[z, x_\gamma] - F'(z)| \leqslant \frac{\gamma}{2} A,$$

we find

$$\|z - (w + v)\| \leqslant \|z - w\| + \|v\|$$
$$\leqslant \eta + \frac{\beta\gamma^2}{2} \|A\|.$$

Hence we will have $K(x_\gamma) \subseteq x_\gamma$ if

$$\eta + \tfrac{1}{2}\beta\gamma^2 \|A\| \leqslant \gamma. \tag{3.7}$$

Since (3.5) holds the inequality (3.7) may be solved to give (3.6). $\blacksquare$

Now we are in a position to compare the two theorems. Take $C = F'(z)^{-1}$, so that $w = p(z)$, $\eta_0 = \eta$, $\beta_0 = \beta$. Thus, for

$$x_\gamma := [z - \gamma e, z + \gamma e] \subset \Omega,$$

where γ is the lower bound in (3.2), one has

$$K(x_\gamma) = w + \gamma(CF[z, x_\gamma] - I)[-e, e].$$

3.3. COROLLARY. *Under the Hypothesis* (K) *with $\chi = \|A\|$, the inequality $h_0 = h \leqslant \frac{1}{2}$ implies $K(x_\gamma) \subseteq x_\gamma$.*

Thus Corollary 3.3 shows that the Moore test (2.5) with the Krawczyk operator is at least as good as and sometimes (as will be shown by the following example) better than the existence test of Kantorivich, both with regard to sensitivity and precision. And of course, it is always better than the test with the derivative in place of the slope. Moreover, it is more versatile in finite precision arithmetic since C can be chosen as an approximate inverse, while an exact inverse is needed for Kantorovich's theorem.

In [9], the example

$$f(\xi) := \xi^2 - a, \qquad 0 \leqslant a \leqslant 1, \quad \xi \in x := [a, 2 - a], \quad z = 1,$$

is used to illustrate that the Moore theorem can be weaker than the Kantorovich theorem. But for this example, the improved Krawczyk operator (2.5) yields an existence condition which is equivalent to the Kantorovich condition. Indeed, in this case,

$$K(x) = \left[\frac{2a-a^2}{2}, \frac{2a+a^2}{2}\right]$$

and $K(x) \subseteq x$ if and only if $0 \leqslant a \leqslant 1$, that is,

$$0 \leqslant h \leqslant \tfrac{1}{2}.$$

As another example we consider

$$f(\xi) := \xi^3 - a, \qquad -1 \leqslant a \leqslant 1, \quad \xi \in x := [a, 2-a], \quad z = 1.$$

With these data we have $w = (2+a)/3$ in Kantorovich's theorem, and

$$h = \tfrac{2}{3}(1-a)(2-a) > \tfrac{1}{2}$$

for all $a \in [0, \frac{1}{2})$.

On the other hand, we have

$$K(x) = \tfrac{1}{3}[a^3 - 6a^2 + 10a + 2, \; -a^3 + 3a^2 - 8a + 6],$$

so that $K(x) \subseteq x$ if $a \in [0.44, \frac{1}{2})$.

Remarks. (1) Concerning the computational complexity, Krawczyk and Neumaier [3] and Neumaier [7] propose a recursive method for calculating a slope $F[z, x]$ of F satisfying (2.2) and hence Theorem 2.1 when F is defined by arithmetic expressions. With those computational rules, the evaluation cost is proportional to the number of operations involved in F. Therefore, it has rather low complexity, in particular lower than when the derivative is computed recursively.

(2) A more complicated existence test, which is also at least as sharp as the Kantorovich test and also uses only first derivative information, has been given by Gay [1, Sect. 3].

References

1. D. M. Gay, Perturbation bounds for nonlinear equations, *SIAM J. Numer. Anal.* **18** (1981), 654–663.
2. L. B. Kantorovich, Functional analysis and applied mathematics, *Uspekhi Mat. Nauk* **3** (1948), 89–185 (in Russian). Translated by C. D. Benster, National Bureau of Standards Report 1509, Washington, DC, 1952.
3. R. Krawczyk and A. Neumaier, Interval slopes for rational functions and associated centered forms, *SIAM J. Numer. Anal.* **22** (1985), 604–616.

4. R. E. Moore, A test for existence of solutions to nonlinear systems, *SIAM J. Numer. Anal.* **14** (1977), 611–615.
5. R. E. Moore, A computational test for convergence of iterative methods for nonlinear systems, *SIAM J. Numer. Anal.* **15** (1978), 1194–1196.
6. R. E. Moore, "Methods and Applications of Interval Analysis," SIAM, Philadelphia, 1979.
7. A. Neumaier, Existence of solutions of piecewise differentiable systems of equations, *Arch. Math.* **47** (1986), 443–447.
8. L. Qi, A generalization of the Krawczyk–Moore algorithm, *in* "Interval Mathematics" (K. Nickel, Ed.), pp. 481–488, New York, 1980.
9. L. B. Rall, A comparison of the existence theorems of Kantorovich and Moore, *SIAM J. Numer. Anal.* **17** (1980), 148–161.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium