Computing [Suppl] 15, 21–28 (2001)

Computing

On the Existence Theorems of Kantorovich, Moore and Miranda

G. E. Alefeld, Karlsruhe, **F. A. Potra**, Baltimore, and **Z. Shen**, Nanjing

Dedicated to Professor Tetsuro Yamamoto on the occasion of his 65th birthday

Abstract

We show that the assumptions of the well-known Kantorovich theorem imply the assumptions of Miranda's theorem, but not vice versa.

AMS Subject Classifications: 65H10.

Key Words: Nonlinear systems, existence theorems, interval analysis.

1. Introduction

The purpose of this paper is to show that the assumptions of the well-known Kantorovich theorem imply the assumptions of Miranda's theorem. This surprising fact can be concluded from a series of well-known results on the existence of solutions of a nonlinear system of equations in $\mathbf{R}^n$. At the end of the paper we give a short direct proof of this fact.

2. Some Existence Theorems and their Relations

Proving the existence of a solution of a system of nonlinear equations is a fundamental problem in nonlinear analysis. In what follows we will review three well-known theorems related to this problem: The Kantorovich Theorem, Moore's Theorem and Miranda's Theorem. The most famous of them is perhaps the Kantorovich Theorem. We will present it both in its "classical" form (see [8] or [10]), and in its "affine-invariant" form proposed by Deuflhard and Heindl [2]. Although the above mentioned theorems hold in general Banach spaces, for the purpose of this paper we will state them only for the n-dimensional space $\mathbf{R}^n$ endowed with the infinity norm. Also, we will present only the parts of the conclusion of those theorems that are relevant to our paper.

Theorem 1. (*Kantorovich*) *Let $f : D \subset \mathbf{R}^n \to \mathbf{R}^n$ be Frechét differentiable in the open convex set D. Assume that for the point $x^0 \in D$ the Jacobian $f'(x^0)$ is invertible with $\|f'(x^0)^{-1}\|_\infty \leq \beta$. Let there be a Lipschitz constant κ* for *f' such that*

$$\|f'(u) - f'(v)\|_\infty \leq \kappa\|u - v\|_\infty \quad \textit{for all } u, v \in D.$$

Let

$$||x^1 - x^0||_\infty = ||f'(x^0)^{-1} f(x^0)||_\infty \le \eta.$$

If $h = \beta\kappa\eta \le \frac{1}{2}$ *and* $\bar{B}(x^0, \rho_-) = \{x \in \mathbf{R}^n \mid ||x - x^0||_\infty \le \rho_-\} \subset D$ *where*

$$\rho_- = \frac{1 - \sqrt{1 - 2h}}{h}\eta,$$

then f *has a zero* $x^* \in \bar{B}(x^0, \rho_-)$. □

Theorem 2. (*Affine-invariant form of the Kantorovich Theorem*) *Let* $f : D \subset \mathbf{R}^n \to \mathbf{R}^n$ *be Frechét differentiable in the open convex set D. Assume that for the point* $x^0 \in D$ *the Jacobian* $f'(x^0)$ *is invertible with*

$$||f'(x^0)^{-1} f(x^0)||_\infty \le \eta. \tag{1}$$

Let there be a Lipschitz constant ω for $f'(x^0)^{-1} f'$ *such that*

$$||f'(x^0)^{-1}(f'(u) - f'(v))||_\infty \le \omega||u - v||_\infty \quad \text{for all } u, v \in D. \tag{2}$$

If $h = \eta\omega \le \frac{1}{2}$ *and* $\bar{B}(x^0, \rho_-) = \{x \in \mathbf{R}^n \mid ||x - x^0||_\infty \le \rho_-\} \subset D$, *where*

$$\rho_- = \frac{1 - \sqrt{1 - 2h}}{\omega}, \tag{3}$$

then f *has a zero* x^* *in* $\bar{B}(x^0, \rho_-)$. □

If the hypothesis of Theorem 1 is satisfied for some constants β, κ and η, then the hypothesis of Theorem 2 is obviously satisfied with $\omega = \beta\kappa$ and η. In many cases the Lipschitz constant ω from Theorem 2 is much smaller than the product of the constants β and κ appearing in Theorem 1. In fact there are examples where Theorem 2 applies but Theorem 1 does not. For a recent application of the affine-invariant form of the Kantorovich Theorem see [9]. The next theorem contains the so-called Moore test [7] which is based on the Krawczyk operator [3].

Theorem 3. (*Moore*) *Let* $f : D \subset \mathbf{R}^n \to \mathbf{R}^n$ *be continuous on the open convex set D. Suppose that for* $x^0 \in D$ *and some real vector* $d = (d_i)$ *with* $d_i \ge 0,\ i = 1, \ldots, n$, *the interval vector* $[x] = [x^0 - d, x^0 + d]$ *is contained in D and*

$$f(x) - f(x^0) = \delta f(x^0, x)(x - x^0), \quad x \in [x],$$

where the matrix $\delta f(x^0, x)$ *is called slope of* f. *Assume furthermore that for some interval matrix* $\delta f(x^0, [x])$ *we have*

$$\delta f(x^0, x) \in \delta f(x^0, [x]) \quad \text{for all } x \in [x]. \tag{4}$$

If there is a real nonsingular matrix A such that the Krawczyk operator

$$K([x],x^0,A) := x^0 - Af(x) + (I - A\delta f(x^0,[x]))([x] - x^0)$$

satisfies

$$K([x],x^0,A) \subset [x], \tag{5}$$

then $[x]$ *contains a zero* x^* *of* f. □

There are different possibilities for finding an interval matrix $\delta f(x^0,[x])$ for which (4) holds, such as the interval arithmetic evaluation of the Jacobian, the interval extension of the Jacobian, the interval extension of the slope etc.

In [11] Rall took for $\delta f(x^0,[x])$ the interval extension of the Jacobian of f on $[x]$ and under this assumption he performed a careful comparison of Theorems 1 and 3 coming to the conclusion that Theorem 1 is more general. Furthermore a simple example is given which shows that (5) does not hold for the ball $\bar{B}(x^0,\rho_-)$ constructed using Theorem 1. However, Theorem 3 is easier to apply. For example, no Lipschitz constant is needed for the Moore test.

Neumaier and Shen [7] took for $\delta f(x^0,[x])$ the interval extension of the slope. In contrast to the preceding case they could show that the Moore test (Theorem 3) always works if Theorem 1 can be applied, but not vice versa. Shen and Wolfe [12] performed the same comparison for Theorem 2 and the Moore test (Theorem 3). The result is the same as in [7].

The last existence theorem to be presented in this paper is Miranda's theorem (see [4] and [6]).

Theorem 4. (*Miranda*) *Let* $f : D \subset \mathbf{R}^n \to \mathbf{R}^n$ *be a continuous function. Assume that the interval vector* $[x] = [x^0 - d, x^0 + d], x^0 = (x_i^0), d = (d_i), d_i \geq 0, i = 1,\dots,n,$ *is contained in D. Let*

$$[x]_i^+ = \{x \in [x], x_i = x_i^0 + d_i\},$$
$$[x]_i^- = \{x \in [x], x_i = x_i^0 - d_i\},$$

be the n pairs of parallel, opposite faces of the interval vector $[x]$. *If*

$$f_i(x)f_i(y) \leq 0 \quad \textit{for all } x \in [x]_i^+, y \in [x]_i^-, \quad i = 1,2,\dots,n,$$

then f has at least one zero x^* *in* $[x]$. □

Very recently Alefeld and Shen [1] proved that if $f(x)$ satisfies the hypothesis of Theorem 3 then $g(x) = Af(x)$ satisfies the hypothesis of Theorem 4, but not the other way around. Their result holds for any interval matrix $\delta f(x^0,[x])$ satisfying (4) independently of how it is obtained (interval arithmetic evaluation of the Jacobian, interval extension of the Jacobian, interval extension of the slope, etc.) Summarizing the above discussion we obtain the following result.

Theorem 5. *Assume either that the assumptions of Theorem 1 or those of Theorem 2 hold for* $f(x)$. *Then the assumptions of Theorem 4 hold for* $g(x) = f'(x^0)^{-1}f(x)$, *but not vice versa.*

This result is quite interesting but the way it was obtained above involves many concepts from interval arithmetic that do not have anything to do with the statement of Theorem 5. It is therefore useful to have a direct proof of this theorem. This will be done in the next section.

3. A Direct Proof

Proof of Theorem 5:

"$\Rightarrow$" As noted in the preceding section, if the hypothesis of Theorem 1 is satisfied then the hypothesis of Theorem 2 is satisfied with $\omega = \beta\kappa$. Therefore in what follows we assume that the hypothesis of Theorem 2 is satisfied. Let e^i denote the i-th unit vector. If one considers for a fixed $i, 1 \le i \le n$, all vectors $u^i = (u^i_j)$ with $(e^i)^T u^i = 1$ and $|u^i_j| \le 1, j = 1, \ldots, n, j \ne i$, then the pairs of opposite faces of the interval vector $[x] = [x^0 - \rho e, x^0 + \rho e], \rho \ge 0, e = (1, \ldots, 1)^T$, can be written as

$$\begin{aligned}[x]^+_i &= \{x \in [x] \mid x_i = x^0_i + \rho\} \\ &= \{x = x^0 + \rho u^i \mid (e^i)^T u^i = 1, |u^i_j| \le 1, j = 1, \ldots, n, j \ne i\}\end{aligned}$$

and

$$\begin{aligned}[x]^-_i &= \{x \in [x] \mid x_i = x^0_i - \rho\} \\ &= \{x = x^0 - \rho u^i \mid (e^i)^T u^i = 1, |u^i_j| \le 1, j = 1, \ldots, n, j \ne i\},\end{aligned}$$

respectively. Let $x = x^0 \pm \rho_- u^i \in [x]^\pm_i$ where ρ_- is defined by (3). We first show that

$$\left| \int_0^1 (e^i)^T \{g'(x^0 \pm t\rho_- u^i) - g'(x^0)\} u^i dt \right| \le \frac{1}{2}\omega\rho_-, \tag{6}$$

where $g(x) = f'(x^0)^{-1}f(x)$. We have

$$\begin{aligned}&\left| \int_0^1 (e^i)^T \{g'(x^0 \pm t\rho_- u^i) - g'(x^0)\} u^i dt \right| \\ &\quad \le \int_0^1 (e^i)^T |g'(x^0 \pm t\rho_- u^i) - g'(x^0)| \cdot |u^i| dt\end{aligned}$$

$$\leq \int_0^1 \sum_{j=1}^n |g'_{ij}(x^0 \pm t\rho_- u^i) - g'_{ij}(x^0)| \cdot |u^i_j| dt$$

$$\leq \int_0^1 \sum_{j=1}^n |g'_{ij}(x^0 \pm t\rho_- u^i) - g'_{ij}(x^0)| dt$$

$$\leq \int_0^1 ||g'(x^0 \pm t\rho_- u^i) - g'(x^0)||_\infty dt$$

$$\leq \omega \int_0^1 || \pm t\rho_- u^i||_\infty dt$$

$$\leq \frac{1}{2}\omega\rho_-.$$

In what follows we use the fact that ρ_- defined by (3) satisfies the quadratic equation

$$\frac{1}{2}\omega\rho_-^2 - \rho_- + \eta = 0. \tag{7}$$

Let $x = x^0 \pm \rho_- u^i \in [x]_i^\pm$. By the well-known generalization of the fundamental theorem of calculus we obtain

$$\begin{aligned} g(x^0 \pm \rho_- u^i) &= g(x^0) \pm \rho_- \int_0^1 g'(x^0 \pm t\rho_- u^i) u^i dt \\ &= g(x^0) \pm \rho_- \int_0^1 \{g'(x^0 \pm t\rho_- u^i) - g'(x^0) + g'(x^0)\} u^i dt \\ &= g(x^0) \pm \rho_- g'(x^0) u^i \pm \rho_- \int_0^1 \{g'(x^0 \pm t\rho_- u^i) - g'(x^0)\} u^i dt. \end{aligned}$$

Using $g'(x^0) = I$ and $(e^i)^T u^i = 1$ gives

$$g_i(x^0 \pm \rho_- u^i) = g_i(x^0) \pm \rho_- \pm \rho_- \int_0^1 (e^i)^T \{g'(x^0 \pm t\rho_- u^i) - g'(x^0)\} u^i dt.$$

By taking the + sign and using (1), (2), (6) and (7) we obtain

$$
\begin{aligned}
&g_i(x^0 + \rho\_u^i) \\
&\quad \geq -|g_i(x^0)| + \rho_- - \rho_-\left|\int_0^1 (e^i)^T \{g'(x^0 + t\rho\_u^i) - g'(x^0)\} u^i dt\right| \\
&\quad \geq -\eta + \rho_- - \frac{1}{2}\omega\rho_-^2 = 0.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
&g_i(x^0 - \rho\_u^i) \\
&\quad \leq |g_i(x^0)| - \rho_- - \rho_- \int_0^1 (e^i)^T \{g'(x^0 - t\rho\_u^i) - g'(x^0)\} u^i dt \\
&\quad \leq \eta - \rho_- + \rho_-\left|\int_0^1 (e^i)^T \{g'(x^0 - t\rho\_u^i) - g'(x^0)\} u^i dt\right| \\
&\quad \leq \eta - \rho_- + \frac{1}{2}\omega\rho_-^2 = 0.
\end{aligned}
$$

Hence the hypothesis of Miranda's Theorem is satisfied for $g(x)$ and $d = \rho\_e$.

"$\Leftarrow$" This part of the proof is shown by using a simple counter-example. Let

$$f(x) = -x^3 + 6x^2 - 11x + 6 = -(x-1)(x-2)(x-3)$$

for $x \in D = [x^0 - d, x^0 + d]$ with $x^0 = 0$ and some fixed $d > 0$. From

$$f'(x) = -3x^2 + 12x - 11$$

it follows $f'(x^0) = -11$ and therefore we have

$$
\begin{aligned}
g(x) &= f'(x^0)^{-1} f(x) \\
&= \frac{1}{11}(x^3 - 6x^2 + 11x - 6).
\end{aligned}
$$

Choosing $d = 1$ it holds that

$$
\begin{aligned}
g(x^0 - d) &= g(-1) = -\frac{24}{11}, \\
g(x^0 + d) &= g(1) = 0
\end{aligned}
$$

and Theorem 4 can be applied to $g(x)$ for $D = [-1, 1]$. On the other hand we obtain immediately

$$|f'(x^0)^{-1} f(x^0)| = \frac{6}{11} =: \eta.$$

Furthermore

$$f'(x^0)^{-1} f'(x) = \frac{3}{11}x^2 - \frac{12}{11}x + 1$$

and therefore

$$\begin{aligned} |f'(x^0)^{-1}(f'(u) - f'(v))| &= \left|\frac{3}{11}(u^2 - v^2) - \frac{12}{11}(u - v)\right| \\ &= \left|\frac{3}{11}(u + v) - \frac{12}{11}\right| |u - v|. \end{aligned}$$

If $u, v \in [-d, d]$ then $\frac{3}{11}(u+v) - \frac{12}{11} \in \left[-\frac{6d}{11} - \frac{12}{11}, \frac{6d}{11} - \frac{12}{11}\right]$ and therefore

$$\left|\frac{3}{11}(u + v) - \frac{12}{11}\right| \leq \frac{6}{11}d + \frac{12}{11}.$$

Hence (2) holds with $\omega := \frac{6}{11}d + \frac{12}{11}$.

Since

$$h := \eta\omega = \frac{6}{11}\left(\frac{6}{11}d + \frac{12}{11}\right) = \frac{6}{121}(6d + 12)$$

is bigger than $\frac{1}{2}$ for all positive d we can not apply Theorem 2. □

It is interesting to note that we have proved more than the fact that the hypothesis of the Kantorovich Theorem in the ∞-norm guarantees the applicability of Miranda's Theorem. Actually we have shown that the function g_i has nonpositive values on the "left face" $[x]_i^-$ and nonnegative values on the "right face" $[x]_i^+$ for all $i \in \{1, \ldots, n\}$.

Acknowledgements

The authors are grateful to U. Schäfer for carefully reading the paper. The work of F. A. Poha was supported in part by NSF, grand DMS-9996154. The work of Z. Shen was performed while he was a visiting professor at the University of Karlsruhe. He would like to acknowledge the support by DFG and the University of Karlsruhe.

References

[1] Alefeld, G., Shen, Z.: Miranda's theorem and the verification of linear complementarity problems. Institut für Wissenschaftliches Rechnen und Mathematische Modellbildung. Universität Karlsruhe (2001).

[2] Deuflhard, P., Heindl, G.: Affine invariant convergence theorems for Newton's method and extension to related methods. SIAM J. Numer. Anal. *16*, 1–10 (1980).

[3] Krawczyk, R.: Newton–Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken. Computing *4*, 187–201 (1969).

[4] Miranda, C.: Un' osservatione su un theorema di Brouwer. Bolletino Unione Mathematica Italiana 2 vol *3*, 5–7 (1940).

[5] Moore, R. E.: A Test for existence of solutions to nonlinear systems. SIAM J. Numer. Anal. *14*, 611–615 (1977).

[6] Moore, R. E., Kioustelidis, J. B.: A simple test for accuracy of approximate solutions to nonlinear (or linear) systems. SIAM J. Numer. Anal. *17*, 521–529 (1980).

[7] Neumaier, A., Shen, Z.: The Krawczyk Operator and Kantorovich's Theorem. J. Math. Anal. Appl. *149*, 437–443 (1990).

[8] Ortega, J. M., Rheinboldt, W. C.: Iterative solution of nonlinear equations in several variables. New York London: Academic Press 1970.

[9] Potra, F. A.: A path-following method for linear complementarity problems based on the affine-invariant Kantorovich Theorem. ZIB-Report 00–30, Konrad-Zuse-Zentrum, Berlin, August 2000.

[10] Rall, L. B.: Computational solution of nonlinear operator equations, New York: John Wiley 1969.

[11] Rall, L. B.: A Comparison of the existence theorems of Kantorovich and Moore. SIAM J. Numer. Anal. *17*, 148–161 (1980).

[12] Shen, Z., Wolfe, M. A.: A note on the comparison of the Kantorovich and Moore theorems. Nonlin. Anal. *15*, 229–232 (1990).

G. E. Alefeld
Institut für Angewandte Mathematik
Universität Karlsruhe
Kaiserstraße 12
D-76128 Karlsruhe
Germany
e-mail: goetz.alefeld@mathematik.uni-karlsruhe.de

F. A. Potra
Department of Mathematics and Statistics
University of Maryland
Baltimore County
1000 Hilltop Circle
Baltimore, MD 21250
USA
e-mail: potra@math.umbc.edu

Z. Shen
Department of Mathematics
Nanjing University
Nanjing, 210093
P. R. China
e-mail: shenzuhe@netra.nju.edu.cn