References

1. F. E. Hohn, Elementary Matrix Algebra, 2nd ed., Macmillan, New York, 1964.
2. P. R. Halmos, Finite Dimensional Vector Spaces, Van Nostrand, Princeton, N. J., 1958.

THE NEWTON-KANTOROVICH THEOREM

J. M. ORTEGA, University of Maryland

One of the basic results in numerical analysis is *The Newton-Kantorovich Theorem:* Let X and Y be Banach spaces and $F: D\subset X\to Y$. Suppose that on an open convex set $D_0\subset D$, F is Frechet differentiable and

$$\|F'(x) - F'(y)\| \leqq K\|x - y\|, x, y \in D_0.$$

For some $x_0\in D_0$, assume that $\Gamma_0\equiv [F'(x_0)]^{-1}$ is defined on all of Y and that $h\equiv\beta K\eta\leqq\frac{1}{2}$ where $\|\Gamma_0\|\leqq\beta$ and $\|\Gamma_0 Fx_0\|\leqq\eta$. Set

$$(1) \qquad t^* = \frac{1}{\beta K}(1 - \sqrt{1 - 2h}), \qquad t^{**} = \frac{1}{\beta K}(1 + \sqrt{1 - 2h})$$

and suppose that $S\equiv\{x|\ \|x-x_0\|\leqq t^*\}\subset D_0$. Then the Newton iterates $x_{k+1} = x_k - [F'(x_k)]^{-1}Fx_k$, $k=0, 1, \cdots$, are well defined, lie in S and converge to a solution x^* of $Fx=0$ which is unique in $D_0\cap\{x|\ \|x_0-x\|<t^{**}\}$. Moreover, if $h<\frac{1}{2}$ the order of convergence is at least quadratic.

Kantorovich has given two basically different proofs of this result using recurrence relations [1] or majorant functions [2]. It is the purpose of this note to give a proof which is a modification of the second approach and is, we believe, easier to understand and present. Moreover, the concept of a majorizing sequence and estimates of the type (5) have been extended [3] to give a convergence theory for a large class of iterative processes. The proof will be an easy consequence of the following lemmas which serve to isolate the essential points.

LEMMA 1. *Let* $\{y_k\}$ *be a sequence in* X *and* $\{t_k\}$ *a sequence of nonnegative real numbers such that*

$$(2) \qquad \|y_{k+1} - y_k\| \leqq t_{k+1} - t_k, \qquad k = 0, 1, \cdots$$

and $t_k\to t^* < \infty$. *Then there exists a* $y^*\in X$ *such that* $y_k\to y^*$ *and*

$$(3) \qquad \|y^* - y_k\| \leqq t^* - t_k, \qquad k = 0, 1, \cdots.$$

The proof is immediate from

$$\|y_{k+p} - y_k\| \leqq \sum_{i=1}^{p} \|y_{k+i} - y_{k+i-1}\| \leqq t_{k+p} - t_k \leqq t^* - t_k,$$

which shows that $\{y_k\}$ is a Cauchy sequence. We shall say that $\{t_k\}$ *majorizes* $\{y_k\}$ if (2) holds.

In the following two lemmas the relevant assumptions of the theorem are assumed to hold.

LEMMA 2. *For all* $x \in Q \equiv \{x \mid \|x-x^0\| < 1/\beta K\} \cap D_0$, $[F'(x)]^{-1}$ *is defined on all of* Y *and*

$$\|[F'(x)]^{-1}\| \leqq \beta/(1-\beta K\|x-x_0\|). \tag{4}$$

If x *and* $Nx \equiv x-[F'(x)]^{-1}Fx$ *are in* Q, *then*

$$\|N(Nx)-Nx\| \leqq \frac{1}{2}\,\frac{\beta K\,\|x-Nx\|^2}{1-\beta K\|x_0-Nx\|}\,. \tag{5}$$

Proof: The first statement follows from the well-known Banach lemma (see, e.g., [4, p. 164]). To prove (5) we note that, since $Fx+F'(x)(Nx-x)=0$,

$$\begin{aligned}\|N(Nx)-Nx)\| &= \|[F'(Nx)]^{-1}F(Nx)\| \\ &\leqq \frac{\beta}{1-\beta K\|x_0-Nx\|}\,\|F(Nx)-Fx-F'(x)(Nx-x)\|\end{aligned}$$

and the result follows by use of the mean value theorem (see, e.g., [2]):

$$\begin{aligned}\|Fy-Fx-F'(x)(y-x)\| &= \left\|\int_0^1 [F'(\theta y+(1-\theta)x)-F'(x)](y-x)d\theta\right\| \\ &\leqq \frac{K}{2}\|y-x\|^2.\end{aligned}$$

LEMMA 3. *The Newton sequence* $\{x_k\}$ *is well-defined and is majorized by the sequence defined by*

$$t_{k+1} = t_k - \frac{(\beta K/2)t_k^2 - t_k + \eta}{\beta K t_k - 1}, \qquad k = 0, 1, \cdots, t_0 = 0. \tag{6}$$

Moreover, $t_k \uparrow t^*$, *where* t^* *is defined by* (1).

Proof: We note first that the t_k are simply the Newton iterates for the polynomial $(\beta K/2)t^2 - t + \eta$ with roots t^* and t^{**} and it follows immediately that $t_k \uparrow t^*$. Now assume that $x_1, \cdots, x_k$ exist and $\|x_i - x_{i-1}\| \leqq t_i - t_{i-1}$, $i=1, \cdots, k$; this holds by assumption for $k=1$. Then $\|x_k - x_0\| \leqq t_k - t_0 \leqq t^*$ so that $x_k \in S$. Hence by Lemma 2, x_{k+1} is defined and

$$\begin{aligned}\|x_{k+1}-x_k\| &= \|N(Nx_{k-1})-Nx_{k-1}\| \leqq \frac{\frac{1}{2}\beta K\|x_k-x_{k-1}\|^2}{1-\beta K\|x_k-x_0\|} \\ &\leqq \frac{\frac{1}{2}\beta K(t_k-t_{k-1})^2}{1-\beta K t_k} = t_{k+1}-t_k,\end{aligned} \tag{7}$$

where the last equality is the result of a simple calculation using the definition of t_k.

The proof of the theorem is now immediate. Lemmas 1 and 3 show that there exists an $x^* \in S$ such that $x_k \to x^*$. That x^* is a solution follows in the usual way from

$$\begin{aligned}\|Fx_k\| &= \|F'(x_k)(x_{k+1} - x_k)\| \leqq [\|F'(x_0)\| + \|F'(x_0) - F'(x_k)\|]\|x_k - x_{k+1}\| \\ &\leqq [\|F'(x_0)\| + Kt^*]\|x_k - x_{k+1}\| \to 0\end{aligned}$$

and the continuity of F in S. If $h < \frac{1}{2}$, the roots t^* and t^{**} are distinct and the order of convergence of t_k to t^* is at least quadratic; hence, by (3) the order of convergence of x_k to x^* is at least quadratic. Finally, the uniqueness statement follows as in [2] by consideration of the simplified Newton iteration $x_{k+1} = x_k - [F'(x_0)]^{-1}Fx_k$.

References

1. L. V. Kantorovich, Functional analysis and applied mathematics, Uspehi Mat. Nauk, 3 (1948) 89–185.

2. L. V. Kantorovich and G. P. Akilov, Functional Analysis in Normed Spaces, Moscow, 1959. Translated by D. Brown, Pergamon, New York, 1964.

3. W. C. Rheinboldt, A unified convergence theory for a class of iterative processes, SIAM J. Numer. Anal., 5 (1968) 42–63.

4. A. E. Taylor, Introduction to Functional Analysis, Wiley, New York, 1957.

MATHEMATICAL EDUCATION NOTES

EDITED BY JOHN R. MAYOR, AAAS
COLLABORATING EDITOR: JOHN A. BROWN, University of Delaware

All material for this department should be sent to John G. Harvey, Dept. of Math., University of Wisconsin, Madison, WI 53706.

'PRIME' PEDAGOGICAL SCHEMES

S. I. BROWN, Graduate School of Education, Harvard University

Recently I. A. Barnett made a plea for requiring a course in introductory number theory, not only of all mathematics majors, but of prospective elementary and secondary school teachers as well [1]. After considering the beauty of such concepts as quadratic residue properties, the Prime Number Theorem, Euclid's proof of the infinitude of primes and Dirichlet's result on arithmetic progressions, he asserts that though not all of these results need be *proved* in such a course, they should be *introduced.* In addition to justification on aesthetic grounds, he feels that concepts of divisibility and related notions will help clarify much of algebra and arithmetic.