PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 62, Number 1, January 1977

QUASI-NONEXPANSIVITY AND TWO CLASSICAL METHODS FOR SOLVING NONLINEAR EQUATIONS

ST. MĂRUŞTER

ABSTRACT. Let $F\colon \mathbf{R}_n \to \mathbf{R}_n$ be a vector-valued function and let $J(x)$ denote the corresponding Jacobi matrix. The main result states that the functions $x - J^{-1}(x) \cdot F(x)$ and $x - \lambda J^T(x) \cdot F(x)$, where λ is a certain positive number, are quasi-nonexpansive. This property is used for establishing the convergence of the Newton and the gradient methods in a finite-dimensional space.

Introduction. In [6] F. Tricomi has studied the convergence of the simple iterates $(x_{k+1} = Tx_k)$ to a fixed point of the real continuous function T. The basic Tricomi's assumption is that the function T is quasi-nonexpansive, that is $|Tx - \xi| < |x - \xi|$, where $x \in \mathbf{R}$ and ξ is a fixed point of T. J. B. Diaz and F. T. Metcalf [2] extended Tricomi's result to n-dimensional spaces as well as to general metric spaces. Ample research concerning the quasi-nonexpansive mappings in Banach or Hilbert spaces has also been given by W. V. Petryshyn and T. E. Williamson, Jr. [5].

The condition of quasi-nonexpansivity seems to be most powerful and so the quasi-nonexpansive mappings seem to be most special. The purpose of this note is to show that the mappings used in two classical methods (the Newton and the gradient methods) are quasi-nonexpansive. This property is used for establishing the convergence of the Newton and the gradient methods.

The generalisation of Tricomi's theorem. Let $\mathbf{R}_n$ be a real n-dimensional space with the norm $\|x\| = \max_{i=1,\dots,n} |x^{(i)}|$, where $x^{(i)}$ denotes the i-component of an $x \in \mathbf{R}_n$, and let $\|A\|$ denote the corresponding norm for a real $n \times n$ matrix $A = [a_{ij}]$ $(i,j = 1, \dots, n)$, that is $\|A\| = \max_i \sum_{j=1}^n |a_{ij}|$. Our development is based on the following generalisation of Tricomi's theorem.

THEOREM 1. *Let T be a continuous mapping of $\mathbf{R}_n$ into $\mathbf{R}_n$ and let ξ be a fixed point of T. Suppose that T is quasi-nonexpansive in the closed ball $D(\xi, r) = \{x \in \mathbf{R}_n\colon \|x - \xi\| \leqslant r\}$, that is*

Received by the editors March 11, 1976.

AMS (MOS) subject classifications (1970). Primary 65H10.

Key words and phrases. Quasi-nonexpansive mapping, nonlinear equation, Newton method, gradient method.

$$(1) \qquad \|Tx - \xi\| < \|x - \xi\|, \qquad x \in D(\xi, r), x \neq \xi.$$

Then ξ is the unique fixed point of T in $D(\xi, r)$ and the sequence $\{x_k\}$ generated by $x_{k+1} = Tx_k$ ($x_0 \in D(\xi, r)$) converges to ξ.

PROOF. From (1) it follows that the sequence $\{\|x_k - \xi\|\}$ is monotone decreasing; therefore $\|x_k - \xi\| \to \rho$ as $k \to \infty$. Suppose that $\rho \neq 0$. There exists a subsequence $\{x_{k_j}\}$ of $\{x_k\}$ which converges to an $x^* \in D(\xi, r)$ and $\|x^* - \xi\| = \rho$. We have

$$\lim_{j\to\infty} x_{k_j+1} = \lim_{j\to\infty} Tx_{k_j} = Tx^*,$$

and since $x^* \neq \xi$, it follows that

$$\rho = \lim_{j\to\infty} \|x_{k_j+1} - \xi\| = \|Tx^* - \xi\| < \|x^* - \xi\| = \rho,$$

a contradiction. It follows that $\rho = 0$ and the sequence $\{x_k\}$ converges to ξ. The uniqueness of the fixed point ξ in $D(\xi, r)$ follows from (1) and the proof is finished.

REMARK. Our proof use arguments similar to those used along the lines indicated in [2].

The Newton-Kantorowich method. Let F: $\mathbf{R}_n \to \mathbf{R}_n$ be a nonlinear vector-valued function, f_i ($i = 1, \ldots, n$) its component and let $J(x)$ denote the Jacobi matrix

$$J(x) = [\partial f_i(x)/\partial x^{(j)}] \qquad (i,j = 1, \ldots, n),$$

where $x^{(j)}$ denotes the j-component of $x \in \mathbf{R}_n$.

THEOREM 2. *Suppose that $\xi \in \mathbf{R}_n$ is a solution of $F(x) = 0$ and that F is a twice continuously differentiable vector-valued function in a certain closed ball $D(\xi, r) \subset \mathbf{R}_n$. Further, suppose that the following conditions are verified:*

(i) *for all $x \in D(\xi, r)$ there exists $J^{-1}(x)$ and $\|J^{-1}(x)\| \leqslant B$,*

(ii) $\sum_{k=1}^{n} |\partial^2 f_i(x)/\partial x^{(j)} \partial x^{(k)}| \leqslant C$ *for all $x \in D(\xi, r)$ and $i, j = 1, \ldots, n$,*

(iii) *the constants r, B, C, satisfies the inequality $nrBC < 2$. Then the function $Tx = x - J^{-1}(x) \cdot F(x)$ is quasi-nonexpansive in $D(\xi, r)$.*

PROOF. Let $R(x) = J(x) \cdot (x - \xi) - F(x)$. Using a well-known theorem (see, for example, [1, p. 459]) and the fact that $F(\xi) = 0$, we have

$$\begin{aligned} \|R(x)\| &= \|F(x) - F(\xi) - J(x) \cdot (x - \xi)\| \\ &\leqslant \frac{1}{2} n \|x - \xi\|^2 \max_{i,j} \sum_{k=1}^{n} \left| \frac{\partial^2 f(\zeta)}{\partial x^{(j)} \partial x^{(k)}} \right|, \end{aligned}$$

where $\zeta = \xi + \theta(x - \xi)$ and $0 < \theta < 1$. Obviously, if $x \in D(\xi, r)$ then $\zeta \in D(\xi, r)$, and from (ii) it follows that

$$\|R(x)\| \leqslant \tfrac{1}{2}n\|x-\xi\|^2 C, \qquad x \in D(\xi, r). \tag{2}$$

This inequality and (i), (iii) give

$$\begin{aligned}\|J^{-1}(x) \cdot R(x)\| &\leqslant \|J^{-1}(x)\| \, \|R(x)\| \leqslant \tfrac{1}{2}n\|x-\xi\|^2 BC \\ &\leqslant \tfrac{1}{2}nrBC\|x-\xi\| < \|x-\xi\|,\end{aligned}$$

for all $x \in D(\xi, r)$, $x \neq \xi$. Now we have

$$\begin{aligned}\|Tx - \xi\| &= \|x - \xi - J^{-1}(x) \cdot F(x)\| \\ &= \|x - \xi - J^{-1}(x) \cdot (J(x) \cdot (x-\xi) - R(x))\| \\ &= \|J^{-1}(x) \cdot R(x)\| < \|x-\xi\|,\end{aligned}$$

and so T is quasi-nonexpansive in $D(\xi, r)$. Theorem 2 is proved.

From Theorems 1 and 2 we obtain

COROLLARY 1. *Let the conditions of Theorem 2 be fulfilled. Then ξ is the unique solution of $F(x) = 0$ in $D(\xi, r)$, and the sequence $\{x_k\}$ generated by the Newton process*

$$x_{k+1} = x_k - J^{-1}(x_k) \cdot F(x_k) \qquad (x_0 \in D(\xi, r)), \tag{3}$$

converges to ξ.

REMARK. This corollary is similar to a classical result obtained by L. V. Kantorowich [3] (see also [1, p. 466]) and, for the case $n = 1$, by A. M. Ostrowski [4].

Let $D(\xi, r')$ be a closed ball with $r' \leqslant r$; obviously, conditions (i) and (ii) are verified in $D(\xi, r')$ with the same constants B and C. Thus, if $r' < 2/nBC$, then condition (iii) is also verified and we obtain

COROLLARY 2. *Let F be as in Theorem 2 and let conditions* (i) *and* (ii) *of Theorem 2 be fulfilled. Further, if x_0 is sufficiently close to ξ, then the Newton process* (3) *converges to ξ.*

It is interesting to obtain the bounds for x_0 which guarantee the convergence of the Newton process. We shall illustrate this question with the following example. Let us consider the equation $x^3 - x = 0$ and its solution $\xi = 0$. A simple computation shows that the conditions of Corollary 1 are verified (hence the Newton process converges) if $|x_0 - \xi| = |x_0| < 1/\sqrt{2} \approx 0.70$. It can also be proved (in this special case) that the conditions of Kantorowich's theorem are verified if $|x_0| < \sqrt{(9 - \sqrt{60})/7} \approx 0.42$. On the other hand, elementary geometrical considerations show that the Newton process converges if $|x_0| < \sqrt{3/7} \approx 0.77$ and that this value cannot be improved.

The gradient method. Let $\|\cdot\|_2$ denote the Euclidean norm, $\|x\|_2 = (\sum_{j=1}^{n} (x^{(j)})^2)^{1/2}$, $x \in \mathbf{R}_n$, and let $\|A\|_2$ denote the corresponding norm for a real $n \times n$ matrix $A = [a_{ij}]$ $(i,j = 1, \dots, n)$, that is $\|A\|_2 = \max_i \lambda_i$, where λ_i are the eigenvalues of AA^T. The majorant norm $\|A\|_2 = (\sum_{i,j} a_{ij}^2)^{1/2}$ can be also used.

Let F: $\mathbf{R}_n \to \mathbf{R}_n$ be a nonlinear vector-valued function, f_i $(i = 1, \dots, n)$ its components, $x^{(j)}$ $(j = 1, \dots, n)$ the components of an $x \in \mathbf{R}_n$, and let $J_{(x)}^T$ be the transposed matrix of the Jacobi matrix. In view of $\|x\| \leqslant \|x\|_2 \leqslant \sqrt{n}\,\|x\|$ for all $x \in \mathbf{R}_n$, relation (2) becomes

$$\|R(x)\|_2 \leqslant \tfrac{1}{2} n^{3/2} \|x - \xi\|_2^2 C. \tag{4}$$

THEOREM 3. *Suppose that* $\xi \in \mathbf{R}_n$ *is a solution of* $F(x) = 0$ *and that* F *is a twice continuously differentiable vector-valued function in a certain closed ball* $\mathfrak{D}(\xi, r) = \{x \in \mathbf{R}_n : \|x - \xi\|_2 \leqslant r\}$. *Further, suppose that the following conditions are verified*:

(i) *for all* $x \in \mathfrak{D}(\xi, r)$ *there exists* $J^{-1}(x)$ *and* $\|J^{-1}(x)\|_2 \leqslant B$,

(ii) $\sum_{k=1}^{n} |\partial^2 f_i(x)/\partial x^{(j)} \partial x^{(k)}| \leqslant C$ *for all* $x \in \mathfrak{D}(\xi, r)$ *and* $i, j = 1, \dots, n$,

(iii) *the constants* r, B, C *satisfy the inequality* $n^{3/2} rBC < 2$.

Then the function

$$Tx = x - \lambda J^T(x) \cdot F(x),$$

where $0 < \lambda \leqslant 1/\max_{x \in \mathfrak{D}(\xi,r)} \|J(x)\|_2^2$, *is quasi-nonexpansive in* $\mathfrak{D}(\xi, r)$.

PROOF. From (i), (iii) and (4) we have

$$\begin{aligned} \|J(x) \cdot (x - \xi)\|_2 &\geqslant (\|J^{-1}(x)\|_2)^{-1} \|x - \xi\|_2 \geqslant B^{-1} \|x - \xi\|_2 \\ &> \tfrac{1}{2} n^{3/2} rC \|x - \xi\|_2 \geqslant \tfrac{1}{2} n^{3/2} \|x - \xi\|_2^2 C \\ &\geqslant \|R(x)\|_2, \end{aligned}$$

for all $x \in \mathfrak{D}(\xi, r)$, $x \neq \xi$. Now since $F(x) = J(x) \cdot (x - \xi) - R(x)$ by a simple computation, we obtain

$$\|Tx - \xi\|_2^2 < \|x - \xi\|_2^2 - \lambda(\|F(x)\|_2^2 - \lambda \|J^T(x)\|_2^2 \|F(x)\|_2^2).$$

But $\lambda \|J^T(x)\|_2^2 \leqslant 1$, and so $\|Tx - \xi\| < \|x - \xi\|$ for all $x \in \mathfrak{D}(\xi, r)$, $x \neq \xi$; that is T is quasi-nonexpansive in $\mathfrak{D}(\xi, r)$. Theorem 3 is proved.

From Theorems 1 and 3 we obtain the following corollary concerning the convergence of the gradient method.

COROLLARY 3. *Let the condition of Theorem* 3 *be fulfilled. Then* ξ *is the unique solution of* $F(x) = 0$ *in* $\mathfrak{D}(\xi, r)$, *and the sequence* $\{x_k\}$ *generated by the gradient method*

$$x_{k+1} = x_k - \lambda J^T(x_k) \cdot F(x_k) \qquad (x_0 \in \mathfrak{D}(\xi, r)),$$

where $0 < \lambda \leqslant 1/\max_{x \in \mathfrak{D}(\xi,r)} \|J(x)\|_2^2$, *converges to* ξ.

References

1. B. P. Demidovič and I. A. Maron, *Foundations of numerical analysis*, 2nd ed., Fizmatgiz, Moscow, 1963. (Russian) MR **27** #4335.

2. J. B. Diaz and F. T. Metcalf, *On the set of subsequential limit points of successive approximations*, Trans. Amer. Math. Soc. **135** (1969), 459–485. MR **38** #2644.

3. L. V. Kantorovič, *On Newton's method*, Trudy Mat. Inst. Steklov **28** (1949), 104–144. (Russian) MR **12**, 419.

4. A. M. Ostrowski, *Solution of equations and systems of equations*, Academic Press, New York, 1960. MR **23** #B571.

5. W. V. Petryshyn and T. E. Williamson, Jr., *Strong and weak convergence of the sequence of successive approximations for quasi-nonexpansive mappings*, J. Math. Anal. Appl. **43** (1973), 459–497. MR **48** #4854.

6. F. Tricomi, *Un teorema sulla convergenza delle successioni formate delle successive iterate di una funzione di una variabile reale*, Giorn. Mat. Battaglini **54** (1916), 1–9.

Department of Mathematics, University of Timisoara, 1900 Timisoara, Romania