

2. H. Robbins and S. Monro, *A stochastic approximation method*, Ann. Math. Statist. vol. 22 (1951) pp. 400–407.

3. J. Wolfowitz, *On the stochastic approximation method of Robbins and Monro*, Ann. Math. Statist. vol. 23 (1952) pp. 457–461.

CORNELL UNIVERSITY

---

## MEAN VALUE METHODS IN ITERATION

W. ROBERT MANN

Due largely to the works of Cesàro, Fejér, and Toeplitz, mean value methods have become famous in the summation of divergent series. The purpose of this paper is to show that the same methods can play a somewhat analogous role in the theory of divergent iteration processes. We shall consider iteration from the limited but nevertheless important point of view of an applied mathematician trying to use a method of successive approximations on some boundary value problem which may be either linear or nonlinear.

It is now widely known that the Schauder fixpoint theorem [1] is a powerful method for proving existence theorems. If one wishes to use it to prove that a given problem has a solution, he proceeds by associating with the problem a convex compact set  $E$  in some Banach space, and a continuous transformation  $T$  which carries  $E$  into itself. Schauder's theorem asserts that  $T$  must have at least one fixpoint, say  $p$ , in  $E$ . If  $E$  and  $T$  have been appropriately chosen, it can then usually be shown that any such fixpoint must be a solution of the original problem and conversely. Mathematical literature since about 1935 abounds with illustrations of this technique. We mention here only [2] and [3] which contain the genesis of the present work.

Let us then begin with a convex compact set  $E$  in a Banach space, and a continuous transformation  $T$  carrying  $E$  into itself. The problem which we shall consider is that of constructing in  $E$  a sequence of elements  $\{x_n\}$  that converge to a fixpoint of  $T$ . Ordinarily one starts by choosing more or less arbitrarily an initial point  $x_1$  in  $E$  and then considering the successive iterates  $\{x_n\}$  of  $x_1$  under  $T$ , where

$$(1) \quad x_{n+1} = T(x_n).$$

If this sequence converges, then obviously its limit point is a fixpoint of  $T$  and the problem is solved. But to guarantee convergence one must impose some further restriction on  $T$ , such as, for example, that

---

Presented to the Society, November 29, 1952; received by the editors October 27, 1952.

it be uniformly distance decreasing. Since, however, there is frequently nothing in the original problem to suggest this ad hoc hypothesis, we are challenged to get along without it.

Assume henceforth that the ordinary iteration process defined by (1) fails to converge. We introduce an infinite triangular matrix  $A$ ,

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

whose elements satisfy the following restrictions:

$$(2) \quad \begin{aligned} a_{ij} &\geq 0 && \text{for all } i \text{ and } j, \\ a_{ij} &= 0 && \text{for all } j > i, \\ \sum_{j=1}^i a_{ij} &= 1 && \text{for all } i. \end{aligned}$$

Starting with an arbitrary element  $x_1$  of  $E$ , we then define the following modified iteration process.

$$(3) \quad x_{n+1} = T(v_n)$$

where

$$(4) \quad v_n = \sum_{k=1}^n a_{nk} x_k.$$

The restrictions (2) rule out the possibility of any of the  $v_n$  lying outside  $E$ . This process is determined by the initial point  $x_1$ , the matrix  $A$ , and the transformation  $T$ . It can be denoted briefly by  $(x_1, A, T)$ , and can be regarded as a generalized iteration process because when  $A$  is the identity matrix,  $I$ ,  $(x_1, I, T)$  is just the ordinary iteration process defined by (1).

**THEOREM 1.**<sup>1</sup> *If either of the sequences  $\{x_n\}$  and  $\{v_n\}$  converges, then the other also converges to the same point, and their common limit is a fixpoint of  $T$ .*

**PROOF.** Let  $\lim x_n = p$ . Since  $A$  is a regular matrix [4],  $\lim v_n = p$ .

<sup>1</sup> Theorem 1 and Theorem 2 are generalizations due to the referee of theorems originally submitted by the author. The author should also like to thank the referee for pointing out simplified methods of proof.

Since  $T$  is continuous,  $\lim T(v_n) = T(p)$ . But  $T(v_n) = x_{n+1}$ , so that  $T(p) = p$ . If now we assume that  $\lim v_n = q$ , then  $\lim x_{n+1} = T(q)$  and by the regularity of  $A$ ,  $\lim v_n = T(q)$ . Hence,  $T(q) = q$ .

If neither  $\{x_n\}$  nor  $\{v_n\}$  has a sequential limit point then, since  $E$  is compact, the set of points comprising each sequence will have more than one limit point. Let  $X$  denote the set of limit points of the  $x$ 's and let  $V$  denote the set of limit points of the  $v$ 's.

**THEOREM 2.** *If the matrix  $A$  satisfies the additional requirements that*

$$(5) \quad \begin{aligned} \lim a_{n,n} &= 0, \\ \lim \sum_{k=1}^n |a_{n+1,k} - a_{n,k}| &= 0, \end{aligned}$$

*then  $X$  and  $V$  are closed connected sets.*

**PROOF.**  $V$  is closed and compact and, because of (5),  $\lim (v_{n+1} - v_n) = 0$ . Therefore  $V$  is connected. Since  $T$  is continuous and  $X = T(V)$ ,  $X$  is closed and connected.

**THEOREM 3.**  *$V$  is included in the convex hull of  $X$ .*

**PROOF.** Let  $\bar{X}$  be the convex hull of  $X$ . By Mazur's theorem [5],  $\bar{X}$  is closed. All but a finite number of the  $x_n$  lie in every open set containing  $\bar{X}$  and hence, for all  $n$  sufficiently large,  $v_n$  will lie arbitrarily close to  $X$ . Therefore the limit of every convergent subsequence of  $\{v_n\}$  lies in  $\bar{X}$  and the theorem is proved.

Now let  $\Lambda$  denote the Cesàro matrix, i.e.

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \cdots \\ 1/2 & 1/2 & 0 & 0 \cdots \\ 1/3 & 1/3 & 1/3 & 0 \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 1/n & 1/n & 1/n \cdots 1/n & 0 & 0 \cdots \end{pmatrix}.$$

$\Lambda$  satisfies all the hypotheses relating to matrices which have been used in this paper. Referring to equations (3) and (4) we see that  $(x_1, \Lambda, T)$  denotes the process of starting with an arbitrary point  $x_1$  in  $E$  and applying the formulas

$$x_{n+1} = T(v_n)$$

where

$$v_n = \frac{1}{n} \sum_{k=1}^n x_k.$$

In other words, the  $(k+1)$ st element in the sequence  $\{x_n\}$  is the image under  $T$  of the centroid of the first  $k$  elements. It is easy to show that in this process

$$(6) \quad v_{n+1} - v_n = \frac{T(v_n) - v_n}{n+1}.$$

Consider now the particular case where the Banach space is just the real axis and the convex compact set  $E$  becomes a closed bounded interval. The following special result is obtained.

**THEOREM 4.** *If  $T$  is a continuous function carrying the interval  $a \leq x \leq b$  into itself and having a unique fixpoint,  $p$ , on  $[a, b]$ , then  $(x_1, \Lambda, T)$  converges to  $p$  for all choices of  $x_1$  on  $[a, b]$ .*

**PROOF.** From (6) it is obvious that  $(v_{n+1} - v_n) \rightarrow 0$ . Since  $T(x)$  is continuous and  $p$  is unique it follows that  $T(x) - x > 0$  if  $x < p$  and  $T(x) - x < 0$  for  $x > p$ . Furthermore, for each  $\delta > 0$  there exists an  $\epsilon > 0$  such that  $|T(x) - x| \geq \epsilon$  whenever  $|x - p| \geq \delta$ . Using (6) to write  $v_{n+1}$  in the following form,

$$v_{n+1} = v_1 + \sum_{k=1}^n \frac{T(v_k) - v_k}{k+1},$$

we see immediately from the above considerations that  $\lim v_n = p$ . By Theorem 1,  $\lim x_n = p$  also.

In higher-dimensional spaces no result comparable to Theorem 4 has been obtained, although in many particular problems the generalized iteration process  $(x_1, \Lambda, T)$  can easily be seen to converge where the ordinary iteration process diverges. For example, if  $E$  denotes a circle plus its interior and  $T$  stands for a rotation of  $\pi/4$  radians about the center, the ordinary iteration process would be useless in an attempt to approximate the unique fixpoint. Using the process  $(x_1, \Lambda, T)$ , however, the sequences  $\{x_n\}$  and  $\{v_n\}$  always spiral into the center regardless of how the initial point is chosen. It seems reasonable to hope that one should be able to prove the convergence of generalized iteration methods under hypotheses weaker than those required to imply convergence of the ordinary iterates. Results in this direction would be of interest, for example, in those nonlinear boundary value problems where one does not wish to invoke a Lipschitz condition to guarantee the convergence of successive approximations.

## BIBLIOGRAPHY

1. J. Schauder, *Der Fixpunktsatz in Funktionalraeumen*, *Studia Mathematica* vol. 2 (1930) pp. 171–180.
2. W. R. Mann and F. Wolf, *Heat transfer between solids and gasses under nonlinear boundary conditions*, *Quarterly of Applied Mathematics* vol. 9 (1951) pp. 163–184.
3. J. H. Roberts and W. R. Mann, *On a certain nonlinear integral equation of the Volterra type*, *Pacific Journal of Mathematics* vol. 1 (1951) pp. 431–445.
4. G. H. Hardy, *Divergent series*, Oxford, 1949.
5. S. Mazur, *Ueber die kleinste konvexe Menge die eine gegebene kompakte Menge enthoelt*, *Studia Mathematica* vol. 2 (1930) pp. 7–9.

UNIVERSITY OF NORTH CAROLINA