The Annals of Mathematical Statistics
1968, Vol. 39, No. 1, 266–269

THE ASYMPTOTIC ERROR OF ITERATIONS[1]

BY PETER FRANK

Syracuse University

0. Iterative methods appear in many mathematical investigations. The theorem concerning contraction mappings is a well known fact about iterations; it is useful for existence theorems and for estimates of error in numerical work. This paper investigates iterative processes in the presence of random errors; it is shown that these errors stabilize.

In Section one, a general theorem is formulated and proven for the case the basic space is the real numbers. In Section two, the theorem is generalized to Banach spaces. In Section three, connections with Markov processes and non-linear integral equations are pointed out.

1. In the statement and proof of the following theorem, we use the notation $\|x\|$ instead of $|x|$ to anticipate the generalization to Banach spaces.

THEOREM 1. *Let*

(a) $T(x)$ *be a real function of a real variable satisfying*

$$\|T(x) - T(y)\| < K\,\|x - y\|, \quad \textit{for all } x \textit{ and } y, \textit{ where} \quad K < 1; \tag{1}$$

(b) $[\epsilon_i]_{i=1}^{\infty}$ *be a sequence of independent and identically distributed random variables with* $E(\|\epsilon_i\|) < \infty$;

(c) $[\delta_i]_{i=1}^{\infty}$ *be a sequence of random variables with* $\sum_{i=1}^{n} K^{n-1}\|\delta_i\|$ *converging to zero in probability (or, equivalently, in distribution);*

(d) *the sequence* $[D_n(X)]_{n=0}^{\infty}$ *be defined by the following relations for each random variable* X:

$$\begin{aligned} D_0(X) &= X, \\ D_n(X) &= T[D_{n-1}(X)] + \epsilon_n + \delta_n \quad \textit{for} \quad n \geqq 1. \end{aligned} \tag{2}$$

Under these conditions, the random variables $D_n(X)$ *converge in distribution; the limiting distribution is determined by the function* T *and the common distribution of the random variables* ϵ_i *and thus does not depend on* X *and the random variables* δ_i.

In the proof of the theorem, the following lemma is needed.

LEMMA. *Let* $X_1, \cdots, X_n, \cdots$ *be a sequence of random variables with* $E(|X_i|) \leqq H < \infty$. *Let* K *be a real number with* $|K| < 1$. *Then the series* $\sum_{n=1}^{\infty} K^n X_n$ *converges absolutely almost everywhere.*

PROOF. This follows directly from the inequalities:

$$E(\textstyle\sum_{n=1}^{N}|K^n X_n|) \leqq \sum_{n=1}^{N}|K|^n H \leqq |K|\cdot H/(1 - |K|) \quad \text{for every} \quad N.$$

PROOF OF THE THEOREM. The case with $\delta_n \neq 0$ can be reduced to the case $\delta_n \equiv 0$ as follows: Define $C_n(X)$ by

Received 25 March 1966; revised 19 July 1967.
[1] Supported by the U. S. Army Research Office (Durham) Contract No. DA-31-124-ARO(D)-87.

$$(3)\qquad \begin{aligned} C_0(X) &= X, \\ C_n(X) &= T[C_{n-1}(X)] + \epsilon_n \quad \text{for} \quad n \geqq 1. \end{aligned}$$

It is easy to establish by induction that $\|D_n(X) - C_n(X)\| \leqq \sum_{i=1}^n K^{n-i}\|\delta_i\|$. By (d), $D_n(X) - C_n(X)$ converges to zero in probability. Thus, if we can show that the sequence $C_n(X)$ has a limiting distribution not depending on X, it will follow that $D_n(X)$ has the same limiting distribution and the proof will be completed. The remainder of the proof consists of three steps.

STEP 1. A sequence of random variables $[Z_n(X)]$ is constructed with the property that for each n, $Z_n(X)$ has the same distribution as $C_n(X)$.

STEP 2. A proof that the sequence $[Z_n(X)]$ converges with probability one. (This is not true of the sequence $[C_n(X)]$.)

STEP 3. A proof that $\lim_{n\to\infty} Z_n(X)$, which exists by Step 2, does not depend on X; that is, if X_1 and X_2 are any two random variables then $\lim_{n\to\infty} Z_n(X_1) = \lim_{n\to\infty} Z_n(X_2)$ with probability one.

Steps one and two imply that the sequence $C_n(X)$ converges in distribution; Step three establishes that the limiting distribution does not depend on X.

PROOF OF STEP 1. For each $n \geqq 1$ define

$$(4)\qquad \begin{aligned} Z_{n,0}(X) &= X \quad \text{and} \\ Z_{n,k}(X) &= T[Z_{n,k-1}(X)] + \epsilon_{n-k+1} \quad \text{for} \quad 1 \leqq k \leqq n. \end{aligned}$$

Let $Z_n(X) = Z_{n,n}(X)$. Comparing (4) with (3), it is seen that $Z_n(X)$ is the same as $C_n(X)$ except that $\epsilon_1, \cdots, \epsilon_n$ have been taken in the reverse order. Since the ϵ's are independent and identically distributed, their joint distribution is invariant under such a permutation; thus $Z_n(X)$ has the same distribution as $C_n(X)$.

When the dependence on X is unimportant, the notations Z_n and C_n are used instead of $Z_n(X)$ and $C_n(X)$ respectively.

PROOF OF STEP 2. To show that the sequence Z_n converges with probability one, it is sufficient to show that the series $\sum_{n=2}^\infty (Z_n - Z_{n-1})$ converges with probability one. To show this, the following inequality is first established:

$$\|Z_n - Z_{n-1}\| < K^{n-1}[\|T(X) - X\| + \|\epsilon_n\|].$$

By (4),
$$\begin{aligned} \|Z_n - Z_{n-1}\| &= \|(TZ_{n,n-1} + \epsilon_1) - (TZ_{n-1,n-2} + \epsilon_1)\| \\ &= \|TZ_{n,n-1} - TZ_{n-1,n-2}\| \\ &< K\|Z_{n,n-1} - Z_{n-1,n-2}\|, \qquad \text{by (1)}. \end{aligned}$$

If this procedure is followed $(n - 2)$ times, it is seen that

$$\begin{aligned} \|Z_n - Z_{n-1}\| &< K^{n-2}\|Z_{n,2} - Z_{n-1,1}\| \\ &= K^{n-2}\|(TZ_{n,1} + \epsilon_{n-1}) - (TX + \epsilon_{n-1})\|, \qquad \text{by (4)}, \\ &= K^{n-2}\|TZ_{n,1} - TX\| \\ &< K^{n-1}\|Z_{n,1} - X\|, \qquad \text{by (1)}, \\ &= K^{n-1}\|(TX + \epsilon_n) - X\|, \qquad \text{by (4)}, \\ &< K^{n-1}[\|TX - X\| + \|\epsilon_n\|], \qquad \text{by the triangle inequality.} \end{aligned}$$

The series $\sum_{n=2}^{\infty} K^{n-1}\|TX - X\|$ converges with probability one because it is a geometric series. The series $\sum_{n=2}^{\infty} K^{n-1}\|\epsilon_n\|$ converges by the lemma and assumption (b). The series $\sum_{n=2}^{\infty}(Z_n - Z_{n-1})$ converges with probability one because each term $(Z_n - Z_{n-1})$ is bounded in absolute value (norm) by the term of a series which converges with probability one.

Proof of Step 3. For any two random variables X_1 and X_2,

$$\begin{aligned} \|Z_n(X_1) - Z_n(X_2)\| &= \|TZ_{n,n-1}(X_1) - TZ_{n,n-1}(X_2)\|, \qquad \text{by (4)}, \\ &< K\|Z_{n,n-1}(X_1) - Z_{n,n-1}(X_2)\|, \qquad \text{by (1)}. \end{aligned}$$

By repeating this procedure n times, it is seen that

$$\|Z_n(X_1) - Z_n(X_2)\| < K^n\|X_1 - X_2\|.$$

Thus, $\lim_{n\to\infty} [Z_n(X_1) - Z_n(X_2)] = 0$ everywhere. Q.E.D.

It is interesting to note that it is not assumed that $E(\epsilon_i) = 0$; thus the system has a stable asymptotic behavior even though a bias may be present at each step. The assumption that $E(\|\epsilon_i\|) < \infty$ was only used to insure that $\sum K^i\|\epsilon_i\|$ converge with probability one; this is not a necessary condition for convergence and undoubtedly can be weakened.

2. In this section, Theorem 1 is generalized to the case that T is a transformation from a real Banach space M into itself, and the random variables δ_i and ϵ_i are M-valued random variables (called M-random variables). General definitions and properties of these random variables can be found in E. Mourier [5] and [6]. Let M^* be the dual of M; by definition, if X is an M-random variable and $m^* \varepsilon M^*$ then $m^*(X)$ is a random variable in the ordinary sense. If the following additional definitions are adopted, Theorem 1 is valid.

Definition (a). Two M-random variables X and Y have the same distribution if and only if for all $m^* \varepsilon M^*$, the random variables $m^*(X)$ and $m^*(Y)$ have the same distribution.

Definition (b). A sequence $[X_i]$ of M-random variables converges in distribution if and only if for all $m^* \varepsilon M^*$ the sequence $[m^*(X_i)]$ converges in distribution.

It is well known that these definitions are equivalent to the usual definitions in n-dimensional space; easy proofs are available using characteristic functions.

In the case of a general Banach space, the proofs of Steps one, two and three are identical. Since $[Z_n]$ converges with probability 1, $m^*(Z_n)$ converges with probability 1 for all $m^* \varepsilon M^*$, and thus $m^*(Z_n)$ converges in distribution. This proves the assertion for $C_n(X)$ and the assertion for $D_n(X)$ follows as before.

3. This section is devoted to a few remarks concerning the fact that the random variables $C_n = C_n(X) (n \geqq 0)$ form a Markov process. (Attention is restricted to the case that M is the set of real numbers; this avoids some technicalities.)

Let the common distribution function of the ϵ's be $E(x)$; that is

$$E(x) = \text{Prob}\,[\epsilon_i \leqq x].$$

It is clear that

$$\text{Prob}\,[C_{n+1} \leqq x \mid C_n = y] = E[x - T(y)] \quad \text{for} \quad n \geqq 0.$$

Let $H(x)$ be the distribution function of the limiting distribution; then H satisfies the equation

$$H(x) = \int_{-\infty}^{+\infty} E[x - T(y)]\, dH(y). \tag{5}$$

In the case that the ϵ's have a density function $e(x)$, the limiting distribution has a density function $h(x)$ which satisfies

$$h(x) = \int_{-\infty}^{+\infty} e[x - T(y)]h(y)\, dy. \tag{6}$$

Theorem 1 shows that if $E(x)$ is a distribution function with $\int |x|\, dE(x) < \infty$ and $T(x)$ satisfies (1) then equation (5) has a unique solution which is a distribution function; a similar statement can be made about Equation (6).

REMARKS. 1. Theorem 1 does not seem to be a consequence of the known facts about Markov processes whose state space is the set of real numbers.

2. The only cases in which an explicit solution to (5) is known are for T linear; the limiting distribution may be singular with respect to Lebesgue measure.

The expected value of the limiting distribution may not be the fixed point even when $E\epsilon_i = 0$. For more information about estimates of the first two moments of the limiting distribution, see Frank [1].

3. In the method of "stochastic approximation" the random variables converge in probability (and sometimes with probability one) in contrast to the processes discussed in this paper.

REFERENCES

[1] FRANK, P. (1960). Iterations with errors. Syracuse University Research Report 28.

[2] KAC, M. (1959). *Statistical Independence in Probability, Analysis and Number Theory.* Wiley, New York.

[3] KOLMOGOROV, A. N. and FOMIN, S. V. (1957). *Elements of the Theory of Functions and Functional Analysis* I. Graylock Press.

[4] LOÈVE, M. (1955). *Probability Theory.* Van Nostrand, Princeton.

[5] MOURIER, E. (1953). Elements aleatoires dans un espace de Banach. *Ann. l'Inst. Henri Poincare.* **13** 161–244.

[6] MOURIER, E. (1956). L-random elements and L^*-random elements in Banach spaces. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **2** 231–242.