TWO INTEGRAL INEQUALITIES

BY
J. R. BLUM AND M. REICHAW
Dedicated to the memory of E. Jabotinsky

ABSTRACT

Let (X, S, μ) and (Y, T, ν) be two measure spaces, $K(g) = \int_Y k(x,y)g(y)d\nu(y)$ $\xi^+ = \max(\xi, 0)$, and $\delta(K) = \sup_{x_1, x_2 \in X} \int_Y (k(x_1, y) - k(x_2, y))^+ \, d\nu(y)$. Two integral inequalities (the first of which has a simple geometrical interpretation) involving $\delta(K)$ are proved. Generalizations of theorems about infinite stochastic matrices and convergence of superpositions of sequences of integral operators are obtained.

1. Let (X, S, μ) and (Y, T, ν) be two measure spaces with totally σ-finite measures and let $(X \times Y, S \times T, \lambda)$ be their product.

Let K be an operator given by

$$K(g) = \int_Y k(x,y)g(y)d\nu(y),$$

where the kernel $k(x,y)$ is a real valued function on $X \times Y$ and $g = g(y)$ is a real function on Y such that $K(g)$ exists.

Denote for a real number ξ

$$\xi^+ = \max(\xi, 0) \text{ and } \xi^- = \max(-\xi, 0).$$

Assuming the existence of the following integral, define

$$(*) \qquad \delta(K) = \sup_{x_1, x_2 \in X} \int_Y (k(x_1, y) - k(x_2, y))^+ d\nu(y).$$

In this paper the following two inequalities (1) and (2) are proved and some of their applications are given. The obtained results are related to some results in [5] and [6]. They also generalize and strengthen some results obtained in [1], [2], [3] and [4].

Received February 2, 1970

FIRST INEQUALITY. *Let $k(x,y) \geqq 0$, $E \subset Y$ be a subset of Y and let*

$$\alpha = \int_X f(x)d\mu(x).$$

Then

$$(1) \quad \int_E d\nu(y)\int_X f(x)k(x,y)d\mu(x) \leqq \delta(K)\int_X f^+(x)d\mu(x) + \alpha \inf_{x\in X}\int_E k(x,y)d\nu(y)\,.$$

SECOND INEQUALITY. *Let $X = Y$, $\mu = \nu$ and let $k(x,y) \geqq 0$. Let $\phi(z,y) = \int_X l(z, x)k(x, y)d\mu(x)$ and $(LK)(g) = \phi(g) = \int_X \phi(z,y)g(y)d\mu(y)$. Denote $\alpha(z_1,z_2) = \int_X(l(z_1,x) - l(z_2,x))d\mu(x)$ and $\bar{\alpha} = \sup_{z_1,z_2\in X}\alpha(z_1,z_2)$. Then*

$$(2) \qquad \delta(\phi) = \delta(LK) \leqq \delta(L)\delta(K) + \bar{\alpha} \inf_{x\in X}\int_X k(x,y)d\mu(y).$$

REMARK 1. Let us note that if $X = Y = N$ is the set of natural numbers with measure $\mu = \nu$ given by:

$\mu(E) = \nu(E) =$ the number of elements of $E \subset N$ and if $k = \{k_{ij}\} = \{k(i,j)\}$ is a stochastic matrix (i.e. $k_{ij} \geqq 0$ and $\Sigma_j k_{ij} = 1$ for all $i \in N$), then

$$(**) \qquad \delta(K) = \sup_{i_1,i_2\in N} \sum_j (k_{i_1 j} - k_{i_2 j})^+ \,.$$

This last function is an estimate of the stability of $\{k_{ij}\}$, see [4, p. 778]. It was used in [4] to obtain ergodic theorems for non-homogeneous Markov chains with possible perturbations and a denumerable number of states. The function (*) is a generalization of (**). It enables us to strengthen and to generalize many theorems in the theory of stochastic matrices (in general infinite) in which (**) has been used. In the sequel $\int\int k(x,y)d\lambda(x,y)$ is denoted by

$$\int\int k(x,y)d\mu(x)d\nu(y).$$

2. Before proving inequalities (1) and (2) we give a formula for $\delta(K)$ expressing $\delta(K)$ in a particular (but important) case in some different way. Although this result will not be necessary in the sequel, it seems to be of some interest because it generalizes some known results and it is quite trivial.

A second expression for $\delta(K)$ in case $C = \int_Y k(x,y)d\nu(y) =$ const. *Suppose that $\int_Y k(x_1,y)d\nu(y) = \int_Y k(x_2,y)d\nu(y) = C$ for all $x_1,x_2 \in X$. Then*

$$(3) \qquad \delta(K) = C - \inf_{x_1,x_2\in X}\int_Y \min\,(k(x_1,y),k(x_2,y))d\nu(y)\,.$$

PROOF. By our assumption one has $\int_Y (k(x_1,y)-k(x_2,y))^+ d\nu(y) = \int_Y (k(x_1,y)-k(x_2,y))^- d\nu(y)$ and $2\int_Y (k(x_1,y)-k(x_2,y))^+ d\nu(y) = \int_Y |k(x_1,y)-k(x_2,y)| d\nu(y)$. By $|\xi-\eta|/2 = (\xi+\eta)/2 - \min(\xi,\eta)$ it follows that

$$\int_Y (k(x_1,y)-k(x_2,y))^+ d\nu(y) = \int_Y \frac{|k(x_1,y)-k(x_2,y)|}{2} d\nu(y) =$$

$$= \int_Y \frac{k(x_1,y)+k(x_2,y)}{2} d\nu(y) - \int_Y \min(k(x_1,y),k(x_2,y)) d\nu(y) =$$

$$= C - \int_Y \min(k(x_1,y),k(x_2,y)) d\nu(y).$$

Hence (3) holds.

REMARK 2. In the case of a finite or infinite stochastic matrix $k = \{k(i,j)\}$ one has (see Remark 1) $C = 1$ and

$$\int_Y \min(k(i_1,j),k(i_2,j)) d\nu(j) = \sum_{j=1}^{\infty} \min(k_{i_1 j}, k_{i_2 j}).$$

In this case formula (3) reduces to $\delta(k) = 1 - \gamma(k)$ where

$$\gamma(k) = \inf_{i_1 i_2} \sum_j \min(k_{i_1 j}, k_{i_2 j}).$$

The result $1-\gamma(k) \leqq \delta(k)$ was recently stated in [3]. The connection between $\delta(k)$ and $\gamma(k)$ for stochastic matrices was investigated in [1] and [2]. The result $1-2\gamma(k) \leqq \delta(k) \leqq 1-\gamma(k)$ was proved in formula (1) of [4], p. 778.

PROOF OF THE FIRST INEQUALITY. By $k(x,y) \geqq 0$ and $\alpha = \int_X f^+(x)d\mu(x) + - \int_X f^-(x)d\mu(x)$ and by Fubini's theorem one has

$$\int_E d\nu(y)\int_X f(x)k(x,y)d\mu(x) = \int_X d\mu(x)\int_E f^+(x)k(x,y)d\nu(y) +$$

$$-\int_X d\mu(x)\int_E f^-(x)k(x,y)d\nu(y) \leqq \left(\int_X f^+(x)d\mu(x)\right) \sup_{x\in X}\int_E k(x,y)d\nu(y) +$$

$$- \left(\int_X (f^+(x)\,d\mu(x) - \alpha)\right) \inf_{x\in X}\int_E k(x,y)d\nu(y) =$$

$$= \left(\int_X f^+(x)d\mu(x)\right)\left[\sup_{x\in X}\int_E k(x,y)d\nu(y) - \inf_{x\in X}\int_E k(x,y)d\nu(y)\right] +$$

$$+ \alpha \inf_{x\in X}\int_E k(x,y)d\nu(y).$$

It remains to note that

$$\sup_{x\in X}\int_E k(x,y)d\nu(y) - \inf_{x\in X}\int_E k(x,y)d\nu(y) =$$

$$= \sup_{x_1x_2\in X}\int_E (k(x_1,y)-k(x_2,y))d\nu(y) \leq$$

$$\leqq \sup_{x_1x_2\in X}\int_X (k(x_1,y)-k(x_2,y))^+ d\nu(y) = \delta(K).$$

Thus (1) is proved.

We state the following two consequences (4) and (5) of inequality (1):

(4) If $\alpha = \int_X f(x)d\mu(x) = 0$ or if $\inf_{x\in X}\int_E k(x,y)d\nu(y) = 0$ then

$$\int_E d\nu(y)\int_X f(x)k(x,y)d\mu(x) \leqq \delta(K)\int_X f^+(x)d\mu(x),$$

(5) If $\alpha = \int_X f(x)d\mu(x) = 0$ then $\int_Y d\nu(y)\int_X |f(x)k(x,y)|\,d\mu(x) \leq$

$$\leqq \delta(K)\int_X |f(x)|\,d\mu(x).$$

PROOF. (4) follows trivially from (1). To prove (5) we note that in case $\alpha = 0$ one has $\int_X |f(x)|\,d\mu(x) = 2\int_X f^+(x)d\mu(x)$. Thus denoting $E_1 = \{y;\ y\in Y$ and $\int_X f(x)k(x,y)d\mu(x) \geqq 0\}$ and $E_2 = \{y; y\in Y$ and $\int_X f(x)k(x,y)d\mu(x) \leqq 0\}$ and substituting in (4) first $E = E_1$ and then $E = E_2$ and then adding both sides of the obtained inequalities one gets

$$\int_Y d\nu(y)\int_X |f(x)k(x,y)|\,d\mu(x) \leqq \delta(K)2\int_X f^+(x)d\mu(x) = \delta(K)\int_X |f(x)|\,d\mu(x).$$

Let us note that Lemma 8 in [3] is a trivial consequence of (5). Let us also note that inequality (1) has a simple geometrical meaning in case $f(x) = 1$ (or more general $f(x) \geqq 0$ interpretated as a mass density). Indeed if $X = Y$ is the real line then $\int\int k(x,y)d\mu(x)d\nu(y)$ is the volume (the mass) bounded i.a. by the surface $z = k(x,y)$, and $\mu(X)\cdot\inf_{x\in X}\int k(x,y)d\nu(y)$ is the volume (the mass) obtained by multiplying the "minimal area" of an x-section $\inf_{x\in X}\int k(x,y)d\nu(y)$ by $\mu(X)$. By (1) their difference is $\leqq \delta(K)\mu(X)$.

PROOF OF THE SECOND INEQUALITY. Fix z_1 and z_2 and let $f(x) = l(z_1,x) - l(z_2,x)$. Then by (1) used for $E = X$ one has:

$$\int_X d\mu(y)\int_X (l(z_1,x)-l(z_2,x))k(x,y)d\mu(x) \leqq \delta(K)\int_X (l(z_1,x)-l(z_2,x))^+ d\mu(x) +$$

$$+\ \alpha(z_1,z_2)\inf_{x\in X}\int_X k(x,y)d\mu(y) \leqq \delta(K)\sup_{z_1,z_2\in X}\int_X (l(z_1,x)-l(z_2,x))^+ d\mu(x) +$$

$$+\ \bar{\alpha}\inf_{x\in X}\int_X k(x,y)d\mu(y.)$$

Hence by $k(x,y) \geqq 0$ it follows that

$$\int_X (\phi(z_1,y)-\phi(z_2,y))^+ d\mu(y) = \int_X d\mu(y)\int_X (l(z_1,x)-l(z_2,x))^+ k(x,y)d\mu(x) \leq$$

$$\leqq \delta(K)\delta(L) + \bar{\alpha}\inf_{x\in X}\int_X k(x,y)d\mu(y).$$

Thus

$$\delta(\phi) = \delta(LK) = \sup_{z_1 z_2\in X}\int_X (\phi(z_1,y)-\phi(z_2,y))^+ d\mu(y)$$

$$\leqq \delta(K)\delta(L) + \bar{\alpha}\inf_{x\in X}\int_X k(x,y)d\mu(y).$$

3. In this section an application of inequalities (1) and (2) to a proof of two theorems on the convergence of a sequence of superpositions of some integral operators is given. The results obtained are related to those proved in [4] and in [5]. The idea of the proof may be easily applied to generalize theorem 1 in [4] p. 780 (after a proper generalization of definitions 1 and 2 in [4]) and many other theorems in the theory of stochastic matrices in which the function (**) is used. We assume in the sequel that $X = Y$, $\mu = \nu$ and consider a family F of integral operators K of the form $K(g) = \int_X k(x,y)g(y)d\mu(y)$. Given two operators K_1 and K_2 with corresponding kernels $k_1(x,y)$ and $k_2(x,y)$ we denote by $k_{12} = \int_X k_1(x,z)k_2(z,y)d\mu(z)$ the kernel corresponding to their product K_1K_2. Similarly, given integral operators $K_{i_1},\cdots,K_{i_n}$ with kernels $k_{i_1}(x,y),\cdots,k_{i_n}(x,y)$ we denote the kernel corresponding to their product $K_{i_1}K_{i_2}\cdots K_{i_n}$ by $k_{i_1\cdots i_n}(x,y)$. For an arbitrary function $r(x,y)$ and an operator R given by

$$R(g) = \int_X r(x,y)g(y)d\mu(y)$$

we denote

(6) $$\|R\| = \sup_{x\in X}\int |r(x,y)|d\mu(y).$$

We prove now the following

THEOREM 1. *Let* $\{K_n\}_{n=1,2}\cdots$ *be a sequence of integral operators with corresponding kernels* $\mathrm{k}_n(x,y) \geqq 0 \quad n = 1,2,\cdots$ *such that*

$$\delta(K_n K_{n-1}\cdots K_1) \to 0 \text{ for } n \to \infty \tag{7}$$

and such that for $n \geqq n_0$ *the kernel* $r_{nm}(x,y)$ *of the operator* $K_{n+m}K_{n+m-1}\cdots \cdots K_{n+1} - I$ *where* I *is the identity operator, satisfies* $\int_X r_{nm}(x,y)d\mu(y) = 0$.

Suppose also that there exists a constant M *such that for* $n \geqq n_0$ *and every* m *one has*

$$\| K_{n+m}K_{n+m-1}\cdots K_{n+1} \| \leqq M. \tag{8}$$

Then $\| K_{n+m}K_{n+m-1}\cdots K_{n+1}K_n\cdots K_1 - K_nK_{n-1}\cdots K_1 \| \to 0$ *for* $n \to \infty$ *(and* m *arbitrary).*

PROOF. Denote $r_{nm}(x,z)$ by $r(x,z)$, $K_n\cdots K_1$ by K and $k_{n+1,n,\ldots 1}(z,y)$ by $k(z,y)$. Then $k(z,y) \geqq 0$ and we have

$$\begin{aligned}
&\| K_{n+m}\cdots K_nK_{n+1}\cdots K_1 - K_n\cdots K_1 \| = \\
&= \| (K_{n+m}\cdots K_{n+1} - I)K_n\cdots K_1 \| = \\
&= \sup_{x\in X}\int_X \left| \int_X r(x,z)k(z,y)d\mu(z) \right| d\mu(y) = \sup_{x\in X} L(x) .
\end{aligned}$$

Now

$$\begin{aligned}
L(x) &= \int_X \left| \int_X r(x,z)k(z,y)d\mu(z) \right| d\mu(y) = \\
&= \int_{E_1(x)} d\mu(y)\int_X r(x,z)k(z,y)d\mu(z) - \int_{E_2(x)} d\mu(y)\int_X r(x,z)k(z,y)d\mu(z)
\end{aligned}$$

where

$$E_1(x) = \{y;\ \int_X r(x,z)k(z,y)d\mu(z) \geqq 0\},$$

and

$$E_2(x) = \{y;\ \int_X r(x,z)k(z,y)d\mu(z) \leqq 0\}.$$

By (1) and by $\int_X r(x,z)d\mu(z) = 0$ it follows that

$$L(x) \leqq \delta(K)\, 2\int_X r^+(x,z)d\mu(z) = \delta(K)\int_X |(r(x,z)|\, d\mu(z).$$

Thus $\sup_{x\in X} L(x) \leqq \delta(K)\| R \|$ where $R = K_{n+m}\cdots K_{n+1} - I$. It follows by (7) and (8) that $\sup_{x\in X} L(x) \leqq \delta(K_n\cdots K_1)\ (M+1) \to 0$ for $n \to \infty$. Theorem 1 is proved.

THEOREM 2. *If assumption* (7) *in Theorem* 1 *is replaced by*

$$\prod_{i=1}^{n}\delta(K_i)\to 0 \text{ for } n\to\infty \tag{7'}$$

and if $C_n = \int_X k_n(x_1,y)d\mu(y) = \int_X k_n(x_2,y)d\mu(y)$ *for each* $x_1,x_2 \in X, n = 1,2,\cdots$ *and all other assumptions of Theorem* 1 *are satisfied then*

$$\| K_{n+m}\cdots K_{n+1}K_n\cdots K_1 - K_nK_{n-1}\cdots K_1 \| \to 0 \text{ for } n\to\infty.$$

PROOF. As in the proof of Theorem 1 one obtains

$\sup_{x\in X} L(x) \leqq \delta(K_n\cdots K_1)(M+1)$ and it remains to apply (2) and (7′).

REMARK 3. Theorems 1 and 2 show that under certain assumptions the sequence $\{K_nK_{n-1}\cdots K_1\}_{n=1,2\ldots}$ is a Cauchy sequence (with the norm defined in (δ)). It follows that if the functions $g = g(y)$ on X form a Banach space Z and if the linear mappings $K_n : Z\to Z$ are continuous mappings of Z into itself, then there exists a continuous and linear operator $K: Z\to Z$ such that

$$\| K_n\cdots K_1 - K \| \to 0 \text{ for } n\to\infty.$$

REMARK 4. It is trivially seen that Theorems 1 and 2 can be generalized to operators K with kernels $k(x,y)$ where x and y are vector variables.

REFERENCES

1. R. L. Dobrushin, *Central limit theorem for non stationary Markov chains I, II*, Theory Prob. Applications, I (1956), 65–80 and 329–383 (English translation).

2. J. Hajnal, *Weak ergodicity in nonhomogeneous Markov chains*, Proc. Cambridge Philos. Soc. **54** (1958), 233–246.

3. A. Paz, *Ergodic theorems for infinite probabilistic tables*, Ann. Math. Statistics **41** (1970), 539–550.

4. A. Paz and M. Reichaw, *Ergodic theorems for sequences of infinite stochastic matrices*, Proc. Cambridge Philos. Soc. **63** (1967), 777–784.

5. M. Reichaw, *On the convergence of superpositions of a sequence of operators*, Studia Math. **25** (1965), 343–351.

6. M. Reichaw, *Supplement to my paper "On the convergence of superpositions of a sequence of operators"*, Studia Math. **28** (1966/7), 361.

TECHNION—ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA
AND
UNIVERSITY OF NEW MEXICO, ALBUQUERQUE