PACIFIC JOURNAL OF MATHEMATICS
Vol. 15, No. 1, 1965

DOUBLY STOCHASTIC OPERATORS OBTAINED FROM POSITIVE OPERATORS

CHARLES HOBBY AND RONALD PYKE

A recent result of Sinkhorn [3] states that for any square matrix A of positive elements, there exist diagonal matrices D_1 and D_2 with positive diagonal elements for which $D_1 A D_2$ is doubly stochastic. In the present paper, this result is generalized to a wide class of positive operators as follows.

Let $(\Omega, \mathfrak{A}, \lambda)$ be the product space of two probability measure spaces $(\Omega_i, \mathfrak{A}_i, \lambda_i)$. Let f denote a measurable function on $(\Omega, \mathfrak{A})$ for which there exist constants c,C such that $0 < c \leqq f \leqq C < \infty$. Let K be any nonnegative, two-dimensional real valued continuous function defined on the open unit square, $(0,1) \times (0,1)$, for which the functions $K(u,\cdot)$ and $K(\cdot,v)$ are strictly increasing functions with strict ranges $(0,\infty)$ for each u or v in $(0,1)$. Then there exist functions $h: \Omega_1 \rightarrow E_1$ and $g: \Omega_2 \rightarrow E_1$ such that

$$\int_{\Omega_2} f(x, v)\, K(h(x), g(v))\, d\lambda_2(v) = 1 = \int_{\Omega_1} f(u,y)\, K(h(u), g(y)) d\lambda_1(u)\ ,$$

almost everywhere $-(\lambda)$.

Let $(\Omega, \mathfrak{A}, \lambda)$ be the product space of two probability measure spaces $(\Omega_i, \mathfrak{A}_i, \lambda_i)$. Let f denote a measurable function on $(\Omega, \mathfrak{A})$ for which there exist constants c,C such that

$$0 < c \leqq f \leqq C < \infty\ . \tag{1}$$

Let K be any nonnegative, real valued continuous function defined on the open unit square, $(0,1) \times (0,1)$, for which the functions $K(u,\cdot)$ and $K(\cdot,v)$ are strictly increasing functions with strict ranges $(0, \infty)$ for each u or v in $(0,1)$.

In what follows, h and g will denote measurable, real valued, functions defined on Ω_1, and Ω_2, respectively. Whenever well defined, set

$$\begin{aligned} R(x: h,g) &= \int_{\Omega_2} f(x,v)\, K(h(x), g(v)) d\lambda_2(v) \\ C(y: h,g) &= \int_{\Omega_1} f(u,y)\, K(h(u), g(y)) d\lambda_1(u) \end{aligned} \tag{2}$$

for $(x,y) \in \Omega$.

Received February 13, 1964. The authors research was supported in part by the National Science Foundation, Grant G-25211 and Grant G-19915.

For a fixed choice of h, g we can think of R and C as defining positive operators. The main result of this paper is that R and C can be made doubly stochastic by choosing h and g appropriately. One immediate consequence of this result is a recent theorem of Sinkhorn [3] on doubly stochastic matrices.

THEOREM. *There exist functions* $h: \Omega_1 \to (0,1)$ *and* $g: \Omega_2 \to (0,1)$ *for which*

$$R(x: h,g) = 1 = C(y: h,g) \,, \tag{3}$$

almost everywhere $-(\lambda)$.

Proof. We shall obtain h and g as the limits of two sequences of functions, $\{h_n\}$ and $\{g_n\}$. The h_n and g_n are defined recursively as follows.

Set $h_0(x) = \alpha$ for all $x \in \Omega_1$, where α is any number in (0,1). If h_n has been defined, let g_n be the function defined by the equation $C(y: h_n,g_n) = 1$. That is, $g_n(y)$ is the solution of the equation

$$1 = \int_{\Omega_1} f(x,y)\, K(h_n(x),g_n(y))d\lambda_1(x) \,. \tag{4}$$

This solution exists and is unique since $C(y: h_n,t)$ is a strictly increasing continuous function of t with range $(0,\infty)$. Furthermore, g_n is easily seen to be measurable if h_n is measurable (certainly the case for h_0), since $\{y \in \Omega_2: g_n(y) \leqq t\} = \{y \in \Omega_2: C(y: h_n,t) \geqq 1\}$ and since $C(y: h_n,t)$ is a measurable function of y for each fixed t. By Fubini's theorem

$$\int_{\Omega_1} R(x: h_n, g_n)d\lambda_1(x) = \int_{\Omega_2} C(y: h_n,g_n)d\lambda_2(y) = 1. \tag{5}$$

Thus if $R(x: h_n,g_n) \geqq 1$ for all x in Ω_1, then $R(x: h_n, g_n) = 1$ almost everywhere $-\lambda_1$, and the proof is complete. If for some $x \in \Omega_1$, $R(x: h_n, g_n) < 1$, we define $h_{n+1}(x)$ to be the numbert for t which $R(x: t,g_n) = 1$. The existence and uniqueness of $h_{n+1}(x)$ follow from our assumptions on K. We set $h_{n+1}(x) = h_n(x)$ at every x where $R(x: h_n,g_n) \geqq 1$. Just as for g_n, we see that h_{n+1} is measurable (since g_n is measurable).

Let $A_n = \{x \in \Omega_1 \mid R(x: h_n,g_n) \leqq 1\}$. If for some $n \geqq 0$, $\lambda_1(A_n) = 1$ we stop our iteration since this implies that $R(x: h_n,g_n) = 1$ a.e. $-\lambda_1$, so we can take h_n and g_n to be h and g of the theorem. We shall assume henceforth that $\lambda_1(A_n) < 1$ for every n.

Observe that $h_{n+1}(x) \geqq h_n(x)$ for every x, thus

$$1 = C(y: h_n,g_n) \leqq C(y: h_{n+1},g_n) \,. \tag{6}$$

Consequently $g_{n+1}(y) \leqq g_n(y)$ for every y. It follows from this monotonicity that the limits $h = \lim_{n\to\infty} h_n$ and $g = \lim_{n\to\infty} g_n$ exist. We shall now show that this choice of h and g satisfies the theorem.

By our construction, $\{A_n\}$ is a nondecreasing sequence of sets. Set $A = \lim_{n\to\infty} A_n$. Since $\lambda_1(A_n) < 1$, the complementary set A_n^c is a set of positive measure for each n. For $x \in A_n^c$, $h_n(x) = \alpha$ whence

$$\begin{aligned} 1 \leqq R(x\colon h_n, g_n) &= \int_{\Omega_2} f(x,y)\, K(\alpha, g_n(y)) d\lambda_2(y) \\ &\leqq C \int_{\Omega_2} K(\alpha, g_n(y)) d\lambda_2(y) \,. \end{aligned}$$

This inequality holds for each n, so one may take limits to obtain

$$1 \leqq C \int_{\Omega_2} K(\alpha, g(y)) d\lambda_2(y) \,.$$

Thus there are positive numbers r and σ such that $\lambda_2\{y \in \Omega_2\colon g(y) \geqq r\} > \sigma$. Then for arbitrary n and $x \in A_n$,

$$1 \geqq c \int_{\Omega_2} K(h_n, g_n) d\lambda_2(y) \geqq c\sigma K(h_n(x), r) \,.$$

Hence, by taking limits on n, one obtains $1 \geqq c\sigma K(h(x), r)$ for each $x \in A$. Let t be a number for which $1 = c\sigma K(t,r)$. Then $h(x) \leqq t$ for $x \in A$, and $h(x) = \alpha$ for $x \in A^c$, whence $h(x) \leqq \beta = \max(\alpha, t) < 1$ for all $x \in \Omega_1$. But for all $y \in \Omega_2$ and all n,

$$\begin{aligned} 1 &= \int_{\Omega_1} f(x,y)\, K(h_n(x), g_n(y)) d\lambda_1(x) \\ &\leqq CK(\beta, g_n(y)) \,, \end{aligned}$$

thus $g(y) \geqq \gamma > 0$ where γ satisfies $C^{-1} = K(\beta, \gamma)$.

The import of the above is that the set $\{(h_n(x), g_n(y))\colon (x,y) \in \Omega,\ n \geqq 0\}$ is contained in a compact subset of the interior of $[0,1] \times [0,1]$, on which K is continuous, and hence bounded. Therefore, by the Lebesgue dominated convergence theorem

$$1 = \lim_{n\to\infty} C(y\colon h_n, g_n) = \int_{\Omega_1} f(x,y)\, K(h(x), g(y)) d\lambda_1(x)$$

and

$$1 = \lim_{n\to\infty} R(x\colon h_{n+1}, g_n) = \int_{\Omega_2} f(x,y)\, K(h(x), g(y)) d\lambda_2(y) \,,$$

for $x \in A$. Moreover

$$1 \leqq \lim_{n\to\infty} R(x\colon h_n, g_n) = \int_{\Omega_2} f(x,y)\, K(h(x), g(y)) d\lambda_2(y) \,,$$

for $x \notin A$. But an inequality here on a set of positive λ_1-measure is impossible by (5), thereby completing the proof.

COROLLARY (Sinkhorn [3]). *Let $A = (a_{ij})$ be an m by m matrix of positive elements. There exist diagonal matrices D_1 and D_2 of positive diagonal elements for which the matrix D_1AD_2 is doubly stochastic.*

Proof. In the above theorem let $\Omega_1 = \Omega_2 = \{1,2, \cdots, m\}$ and let $\lambda_1 = \lambda_2$ be the uniform measure, $\lambda_1(\{j\}) = 1/m$. Set $K(u,v) = uv(1-u)^{-1}(1-v)^{-1}$ and $f(i,j) = a_{ij}$. By the theorem there exist functions h and g such that

$$m^{-1}\sum_{i=1}^{m} a_{ij}h(i)g(j)\,[1-h(i)]^{-1}\,[1-g(j)]^{-1} = 1$$
$$= m^{-1}\sum_{j=1}^{m} a_{ij}h(i)g(j)\,[1-h(i)]^{-1}[1-g(j)]^{-1}\,.$$

The corollary is then proved if one lets $d_{1i} = m^{-1/2}[1-h(i)]^{-1}h(i)$ and $d_{2i} = m^{-1/2}[1-g(i)]^{-1}g(i)$ be the diagonal elements of D_1 and D_2 respectively.

The above result for symmetric matrices has also been obtained by Marcus and Newman [1] and Maxfield and Minc [2].

The application which motivated Sinkhorn's theorem was the case in which A is the matrix of maximum likelihood estimates of a stochastic transition matrix P of a Markov Chain. When it is further known that P is actually doubly stochastic, then Sinkhorn's result shows that numbers $\{x_1, \cdots, x_n; y_1, \cdots, y_n\}$ exist such that A can be renormalized by dividing the ith row by x_i and the jth column by y_j to obtain a doubly stochastic matrix. However, if one considers the maximum likelihood equations for the restricted case in which P is known to be doubly stochastic one observes that the proper normalized form of A (relative to the maximum likelihood approach) is a doubly stochastic matrix $B = (b_{ij})$ with $b_{ij} = a_{ij}(x_i + y_j)^{-1}$. The existence of such a normalization follows straightforwardly from the proof of the above theorem. To see this, consider the function $K(u,v) = [v^{-1} - (1-u)^{-1}]^{-1}$ defined on the triangular region $u > 0$, $v > 0$, $u + v < 1$. This function is nonnegative and continuous on this triangle. Moreover, both $K(u,\cdot)$ and $K(\cdot,v)$ are strictly increasing functions wherever defined and the ranges of $K(u,\cdot)$ and $K(\cdot,v)$ are respectively $(0,\infty)$ and $(v[1-v]^{-1},\infty)$ for each fixed u and v. Let λ_1 and λ_2 be the same discrete measures as used in the proof of the above corollary. The functions $R(x\colon h_n,g_n)$ and $C(y\colon h_n,g_n)$ then become finite sums. The only change required in the proof is that one must show that the points $(h_n(x),g_n(y))$, for all $n \geqq 1$ and all x and y, are well defined and contained in a compact subset of the domain of K. That this is true follows from the assumptions on the monotonicity, continuity and range of K, combined with the fact that the integrals are finite sums. Actually, because of these properties, it is clear that $K(h_n(x), g_n(y))$ is bounded by mc^{-1} for all n and y.

REFERENCES

1. M. Marcus and M. Newman. *The permanent of a symmetric matrix*, to appear. [cf. Notices of Amer. Math. Soc. **8** (1961), Abstract 587-85, p. 595.]

2. H. Minc and J. E. Maxfield, *A doubly stochastic matrix equivalent to a given matrix*, Notices of Amer. Math. Soc. **9** (1962), Abstract 592-81, p. 309.

3. R. Sinkhorn, *A relationship between arbitrary positive matrices and doubly stochastic matrices*, Ann. Math. Stat. **35** (1964), 876-879.

UNIVERSITY OF WASHINGTON