PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 79, Number 3, July 1980

BALAYAGE DEFINED BY THE NONNEGATIVE CONVEX FUNCTIONS[1]

P. FISCHER AND J. A. R. HOLBROOK

ABSTRACT. We study the Choquet order induced on measures on a linear space by the cone of nonnegative convex functions. We are concerned mainly with discrete measures, and the following result is typical. Let $x_1, \ldots, x_r, y_1, \ldots, y_n$, where $r < n$, be points in $\mathbf{R}^d$. Then

$$\sum_1^r f(x_k) < \sum_1^n f(y_k)$$

for all nonnegative, continuous, convex functions f if, and only if, there exists a doubly stochastic matrix M such that

$$x_j = \sum_{k=1}^n m_{jk} y_k \qquad (j = 1, \ldots, r).$$

In the case $d = 1$, this result may be found in the work of L. Mirsky; our methods allow us to place such results in a general setting.

1. The key step. Here we deal with Baire measures (always nonnegative and finite) on a compact subset K of a Hausdorff space E. We shall write $\mu(f)$ for the integral of a continuous (real-valued) function f on K with respect to the measure μ, and δ_x for the unit mass at $x \in K$. When K has a convex structure, we write C for the cone of continuous convex functions on K and C^+ for the nonnegative members of C.

If $\mu(f) < \nu(f)$ for all $f \in C^+$, we follow a common terminology (see, e.g., P. A. Meyer [3, Chapter XI, §3]) in saying that the measure ν is a "balayage" of μ relative to the class C^+. The following theorem relates balayage relative to C^+ to balayage relative to C.

THEOREM 1. *Let μ, ν be two Baire measures on a compact convex subset K of a locally convex topological vector space E. If ν is a balayage of μ relative to C^+, then there exists a Baire measure λ such that ν is a balayage of $\mu + \lambda$ relative to C. Furthermore we may choose λ to be concentrated at a point $x_0 \in K$ if we wish*: $\lambda = (\nu(1) - \mu(1))\delta_{x_0}$.

PROOF. By the Riesz representation theorem, each (nonnegative) Baire measure λ corresponds to a continuous linear functional ϕ on $C(K)$ such that $\phi(f) > 0$ for all f in the (convex) cone P of positive functions in $C(K)$. If we find such a ϕ with the additional properties

Received by the editors July 26, 1978 and, in revised form, July 27, 1979.
1980 *Mathematics Subject Classification*. Primary 26B25, 26D15, Secondary 15A39, 15A51.
Key words and phrases. Hardy-Littlewood-Pólya order, doubly stochastic matrices, balayage.
[1]Work supported in part by the National Research Council of Canada under grant A-8745.

0002-9939/80/0000-0320/$02.00

$$\phi(f) \leq (\nu - \mu)(f) \qquad (f \in C^+)$$

and

$$\phi(1) = (\nu - \mu)(1),$$

then ν is a balayage of $\mu + \lambda$ relative to C, since $C = \{f + r1 : f \in C^+, r \in \mathbf{R}\}$.

If $\mu(1) = \nu(1)$ we can simply put $\phi = 0$. Otherwise, setting $\alpha = \nu - \mu$, we have $\alpha(1) > 0$ and we can define the linear operator T on $C(K)$ by

$$Tf = f - (\alpha(f)/\alpha(1))1.$$

Now $f \in C^+$ implies $Tf \in C$ and, we claim, $Tf \notin P$; otherwise we would have $0 < m = \min Tf$, and $Tf - m1 \in C^+$ so that

$$0 \leq \alpha(Tf - m1) = \alpha(f - (\alpha(f)/\alpha(1))1) - \alpha(m1) = -m\alpha(1) < 0,$$

a contradiction. Thus the cone TC^+ and the open cone P do not intersect, so that by a well-known form of the separation theorem (see, e.g., W. Rudin [6, Theorem 3.4]) there is a continuous linear functional $\psi \not\equiv 0$ on $C(K)$ such that $\psi(f) \leq \psi(g)$ whenever $f \in TC^+$ and $g \in P$.

Since P is an open cone and $\psi \not\equiv 0$, it is clear that $\psi(P)$, being bounded below, must be $(0, \infty)$. If we set $\phi = (\alpha(1)/\psi(1))\psi$, then $\phi(1) = \alpha(1)$, and it remains to show that $\phi(f) \leq \alpha(f)$ whenever $f \in C^+$. But in this case we have ensured that $\phi(Tf) \leq 0$, so that $\phi(f) \leq \phi((\alpha(f)/\alpha(1))1) = \alpha(f)$.

Finally we note that we can, if we wish, replace λ by its *resultant* $\lambda(1)\delta_{x_0}$; more precisely, let x_0 be the barycentre of the probability measure $\lambda/\lambda(1)$ (see R. R. Phelps [5, Proposition 1.1]). By definition, x_0 is that point in K such that $\delta_{x_0}(f)$ $(= f(x_0)) = \lambda(f)/\lambda(1)$ for all continuous affine functions f on K, and it is well known that the inequality $\lambda(1)f(x_0) \leq \lambda(f)$ follows for all $f \in C$. Thus we may replace λ by $\lambda(1)\delta_{x_0} = (\nu(1) - \mu(1))\delta_{x_0}$. Q.E.D.

2. Application. As we shall make clear in the remarks below, the following theorem provides a common extension of some basic results on balayage for discrete measures.

THEOREM 2. *Let* $x = (x_1, \ldots, x_r)$, $y = (y_1, \ldots, y_n)$ *where* $r \leq n$ *and* x_k, y_k *are elements of* $\mathbf{R}^d$. *Then the following are equivalent.*

(i) $\Sigma_1^r f(x_k) \leq \Sigma_1^n f(y_k)$ *for every convex continuous function* $f: K \to \mathbf{R}^+$, *where* K *is the convex hull in* $\mathbf{R}^d$ *of the* x's *and* y's.

(ii) $x = [My]_r$ *for some doubly stochastic matrix* M (*here* $[z]_r$ *denotes the vector formed by the first* r *components of the vector* z *and the product* My *is interpreted formally with* y *as a column vector*).

Remarks. (a) In the one-dimensional case ($d = 1$) this result comes from Ch. Davis and L. Mirsky (see [4]). In that case, the well-known relationship between doubly stochastic matrices and the Hardy-Littlewood-Pólya order allows the addition of a third equivalent statement:

(iii) $x = [z]_r$ for some $z \in \mathbf{R}^n$ such that

$$\sum_{k=1}^{i} z_k^* \leq \sum_{k=1}^{i} y_k^*$$

for $i = 1, \ldots, n$ with equality holding for $i = n$, where z^* and y^* denote the nonincreasing rearrangements of z and y.

(b) The general d-dimensional form of Theorem 2 is due to S. Sherman [7] and C. Stein (see D. Blackwell [1]) in the special case $r = n$. Note that in this case there is no need to require nonnegative functions in (i) since the inequality is unchanged upon adding any constant to f. We shall prove Theorem 2 by deriving it from the special case of Sherman and Stein.

PROOF. (i) $\Rightarrow$ (ii). In the terminology of Theorem 1, the hypothesis says that $\nu = \Sigma_1^n \delta_{y_k}$ is a balayage of $\mu = \Sigma_1^r \delta_{x_k}$ relative to C^+. By Theorem 1, there exists x_0 such that ν is a balayage of $\mu + (n - r)\delta_{x_0}$ relative to C. Thus $z = (x_1, \ldots, x_r, x_0, \ldots, x_0)$ and y satisfy the hypothesis (i) of the Sherman-Stein theorem, so that $z = My$ for some doubly stochastic M. Hence $x = [z]_r = [My]_r$.

(ii) $\Rightarrow$ (i). By the theorem of Sherman and Stein and the nonnegativity of f, we have, with $z = My$,

$$\sum_1^n f(y_k) \geqslant \sum_1^n f(z_k) \geqslant \sum_1^r f(z_k) = \sum_1^r f(x_k).$$

Q.E.D.

3. A variant. In [2] an analogue of the Sherman-Stein theorem is established for substochastic matrices. One form of this result may be stated as follows (cf. [2, Théorème 8]).

THEOREM 3. *Let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ where x_k and y_k are elements of $\mathbf{R}^d$. Then the following statements are equivalent.*

(i) *$\Sigma_{k=1}^n f(x_k) \leqslant \Sigma_{k=1}^n f(y_k)$ for every continuous convex function f: $\mathbf{R}^d \to \mathbf{R}$ such that $f \geqslant f(0)$.*

(ii) *there exists an $n \times n$ doubly substochastic matrix M such that $x = My$.*

The following theorem extends Theorem 3 in the same way that Theorem 2 extends the Sherman-Stein theorem.

THEOREM 4. *Let $x = (x_1, \ldots, x_r)$, $y = (y_1, \ldots, y_n)$, where $r \leqslant n$ and the x_k and y_k are elements of $\mathbf{R}^d$. Then the following statements are equivalent.*

(i) *$\Sigma_{k=1}^r f(x_k) \leqslant \Sigma_{k=1}^n f(y_k)$ for every continuous convex function f: $\mathbf{R}^d \to \mathbf{R}^+$ such that $f \geqslant f(0)$.*

(ii) *there exists an $n \times n$ doubly substochastic matrix M such that $x = [My]_r$.*

PROOF. (ii) $\Rightarrow$ (i). This follows by an obvious modification of our proof of the corresponding implication in Theorem 2.

(i) $\Rightarrow$ (ii). For any convex continuous f: $\mathbf{R}^d \to \mathbf{R}$ such that $f \geqslant f(0)$ the function $f - f(0)$ satisfies the hypotheses and the resulting inequality clearly implies that

$$\sum_1^r f(x_1) + (n - r)f(0) \leqslant \sum_1^n f(y_k).$$

Using the implication (i) $\Rightarrow$ (ii) of Theorem 3, we see that there exists a doubly substochastic matrix M so that $(x_1, \ldots, x_r, 0, \ldots, 0) = M(y_1, \ldots, y_n)$. Q.E.D.

References

1. D. Blackwell, *Equivalent comparisons of experiments*, Ann. Math. Statistics **24** (1953), 265–272.

2. P. Fischer and J. A. R. Holbrook, *Matrices sous-stochastiques et fonctions convexes*, Canad. J. Math. **29** (1977), 631–637.

3. P. A. Meyer, *Probability and potentials*, Blaisdell, Waltham, Mass., 1966.

4. L. Mirsky, *Majorization of vectors and inequalities for convex functions*, Monatsh. Math. **65** (1961), 159–169.

5. R. R. Phelps, *Lectures on Choquet's theorem*, Van Nostrand, Princeton, N. J., 1966.

6. W. Rudin, *Functional analysis*, McGraw-Hill, New York, 1973.

7. S. Sherman, *On a theorem of Hardy, Littlewood, Pólya and Blackwell*, Proc. Nat. Acad. Sci. U. S. A. **37** (1951), 826–831.

Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada