Thickness of families of sets and a minimax lemma

Vlastimil Pták

Institute of Mathematics, Czechoslovak Academy of Sciences, 115 67 Prague 1, Czechoslovakia

Received 4 January 1991

Abstract


A dual expression for the thickness of a family of sets is given. The thickness is a numerical characteristic of families of sets introduced in Lešanovský and Pták (1986). The dual expression is based on a simple combinatorial minimax result.

1. Introduction

The combinatorial lemma on the existence of convex means [6] proved by the author in 1959 was intended as a tool for the investigation of some basic questions of mathematical analysis concerning the interchange of the order of two limit operations—in particular questions generalizing classical results like the Lebesgue dominated convergence theorem, the Fubini theorem and related results. The connection between the lemma, the statement of which is purely combinatorial, and analysis is not at all obvious at first glance. The author showed, however, that the lemma indeed represents the combinatorial essence of a number of results in what could be described in today’s terminology as weak compactness. To explain the motivation for the study of convex means it suffices to recall the fact that the weak closure of a set $M$ in a Banach space $E$ may be characterized as the set of those elements of $E$ that may be arbitrarily well approximated in the uniform topology by convex combinations of points in $M$. Trying to isolate the combinatorial substance of the results on weak compactness the author realized that the relevant combinatorial problem may be formulated as a problem of distributing a unit mass into a finite number of points in a certain manner. More precisely, the combinatorial problems to be considered are of the following type.

We are given a set $S$ and a family $W$ of subsets of $S$. Furthermore, a small positive number $\epsilon$, a safety margin, is prescribed. Our task consists in dividing a
unit mass into a finite number of parts to be situated at certain points in $S$ in such a manner that no set $w$ of the family $W$ contains total weight exceeding $\varepsilon$.

What is the lowest possible safety margin that can be imposed on a given family $W$? We shall call it the thickness of the family and denote it by $e(W)$. It is given by the formula

$$e(W) = \inf_{\lambda \in P(S)} \sup_{w \in W} \lambda(w).$$

Here $P(S)$ is the set of all formal convex combinations of points in $S$, in other words all nonnegative functions $\lambda$ with finite support defined on $S$ such that $\sum_{s \in S} \lambda(s) = 1$. The elements of $P(S)$ may be considered as probability measures on $S$ and $\lambda(w)$ is taken to mean

$$\lambda(w) = \sum_{s \in w} \lambda(s).$$

Obviously the safety margin considered above may not be pushed under this number $e(W)$; this justifies the name given to this numerical characteristic of the family $W$. It is to be expected that in applications to approximation problems families $W$ with $e(W) = 0$ will be of particular importance; in fact, somewhat stronger postulates have to be imposed on the family $W$: for applications to functional analysis it was necessary to characterize those families $W$ for which $e(W) = 0$ in a stronger sense, hereditarily. More precisely, it was necessary to characterize the families $W$ such that $e(W \cap R) = 0$ for each infinite $R \subset S$. To explain the notation: $W \cap R$ stands for the family of all intersections $w \cap R$, $w \in W$. The characterization, the main result of [6] may be reformulated as follows: if $e(W \cap R) > 0$ for some infinite $R$ then there exists an infinite $N$ for which $e(W \cap N) = 1$, in the language of combinatorics:

*for every finite $F \subset N$ there exists a $w \in W$ such that $F \subset w$.*

The problem of characterizing families $W$ for which $e(W) = 0$ remained open. We intend to present, in the present note, a dual description of the characteristic $e(W)$ which may be used, in particular, to give a solution of this problem.

Duality plays an important role here as it did in the earlier stages of the theory. In fact, the method used by the author for the study of weak compactness was based on establishing complete duality: treating a family $F$ of functions on a set $T$ as a function of two variables $B(f, t)$ defined on $F \times T$ by the formula

$$B(f, t) = f(t).$$

Extending $B$ to a bilinear form in the obvious manner it is possible to consider not only convex combinations of functions but also convex combinations of points in $T$: they act as convex combinations of Dirac measures. If we take, on $F$, the topology of pointwise convergence and on $T$ the weak topology generated by $F$, $B$ will be a separately continuous bilinear form. Criteria of weak compactness
may be obtained by establishing conditions under which $B$ may be extended to the compactifications of $P(F)$ and $P(T)$.

In conformity with this general principle we intend to treat families $W$ of subsets of a set $S$ as a relation $R$ on $W \times S$; the relation $R$ is defined by the requirement $[w, s] \in R$ iff $s \in w$. In this manner the family $W$ appears as the family of all sections of $R$. To each $w \in W$ we assign the subset of $s$

$$w \rightarrow \{ s \in S, [w, s] \in R \}.$$ 

Clearly every family of subsets of $S$ may be obtained as the family of sections of a suitable relation. We shall use the same letter $R$ for the characteristic function of the relation $R$ so that $R(w, s) = 1$ iff $s \in w$ and $R(w, s) = 0$ otherwise.

The thickness of a family $W$ of subsets of $S$ is defined as

$$\inf_{\lambda \in P(S)} \sup_{w \in W} \lambda(w) = \inf_{\lambda \in P(S)} \sup_{w \in W} \sum \lambda_j R(w, s_j).$$

The point of view based on duality makes it possible to view the sets $w$ as functions on $S$ by identifying them with the corresponding characteristic functions. Accordingly, convex combinations of the sets $w$ are meaningful and their use makes it possible to see the notion of thickness in a different light.

2. The dual expression

We present another equivalent expression of the characteristic $e(W)$ which gives further support to the intuitive interpretation of $e(W)$ as thickness.

We start by considering the families investigated in the paper [1].

(1) Consider the case $S = \{1, 2, \ldots, n\}$ and take for $W$ the family of all subsets of cardinality $k$. Identifying the sets $w \in W$ we may consider the mean

$$b = \left( \binom{n}{k}^{-1} \right) \sum_{w \in W} w.$$ 

Given $s \in S$ there are exactly $\binom{n-1}{k-1}$ sets $w \in W$ such that $s \in w$. It follows that

$$b(s) = \left( \binom{n}{k}^{-1} \binom{n-1}{k-1} \right) = \frac{k}{n}$$ 

for every $s$.

We have thus found a function $b \in \text{conv } W$ with $b(s) = k/n$ for every $s \in S$.

(2) In [1] we have described, for each irrational $\alpha$ between zero and one, a family $W$ for which $e(W) = \alpha$. The set $S$ was the set of all rational numbers of the form $r/2^n$ with arbitrary $n$ and $0 \leq r \leq 2^n$. The family $W$ was the union of subfamilies $W_n$

$$W_n = \{ w_{nj}, 0 \leq j < 2^n \},$$

$$w_{nj} = \left\{ \frac{j}{2^n}, \frac{j+1}{2^n}, \ldots, \frac{j+k_n-1}{2^n} \right\},$$

where $k_n$ is a suitable integer for each $n$. We now see that $e(W_n) = k_n/n$. The reason is that $\inf_{\lambda \in P(S)} \sup_{w \in W} \lambda(w) = \inf_{\lambda \in P(S)} \sup_{w \in W} \sum \lambda_j R(w, s_j) = 1/n.$

The family $W$ then gives another interpretation of the thickness $e(W)$. It is the family of all sections of $R$.
the numerators being taken modulo $2^n$. The numbers $k_n$ are determined by the requirement that

$$k_n - 1 < 2^n \alpha \leq k_n;$$

they describe the approximation of $\alpha$ by rational numbers of denominator $2^n$. Given any finite set $F \subset S$, there exists an integer $n$ such that $2^n F$ consists of integers only. Thus

$$F \cap F_n = \left\{ \frac{r}{2^n} : 0 \leq r < 2^n \right\}.$$  

Set

$$b = 2^{-n} \sum_{w \in W} w$$

so that $b(s) = k_n/2^n$ for all $s \in F_n$. Since $0 \leq \alpha - k_n/2^n < 1/2^n$ we have $|b(s) - \alpha| < 1/2^n$ for all $s \in F_n$.

In the first case, the inequality $e(W) \geq k/n$ was a consequence of the fact that $\text{conv } W$ contained a function $b$ with $b(s) \geq k/n$ for all $s \in S$.

In the second case we could prove, for every finite $F \subset S$ and every $\epsilon > 0$, the existence of a $b \in \text{conv } W$ for which $b(s) \geq \alpha - \epsilon$ for all $s \in S$.

It is easy to see that for $\alpha > 0$, the existence of such a $b$ implies the inequality $e(W) \geq \alpha$; we intend to show that the inequality $e(W) \geq \alpha$ is in fact equivalent to the existence of such a $b$. Indeed, this is a consequence of the following obvious lemma.

**Lemma.** Let $\epsilon > 0$ and suppose that, for each finite $F \subset S$ and each $\epsilon' < \epsilon$ there exists a $b \in \text{conv } W$ such that $b(s) \geq \epsilon'$ for all $s \in F$. Then $e(W) \geq \epsilon$.

**Proof.** Suppose, on the contrary, that $e(W) < \epsilon$. Then there exists a $\lambda \in P(S)$ such that

$$\sigma = \sup_{w \in W} \lambda(w) < \epsilon.$$

Let $F$ be the carrier of $\lambda$. By assumption, there exists a $b \in \text{conv } W$ such that $b(s) \geq \frac{1}{2}(\sigma + \epsilon)$ for all $s \in F$. It follows that

$$\frac{1}{2}(\sigma + \epsilon) \leq \lambda(b) \leq \sup_{w \in W} \lambda(w) = \sigma$$

and this is a contradiction. □

Denote by $K$ the closed segment $[0,1]$ and consider the cartesian product $K^S$; clearly the assumption of the preceding lemma is equivalent to the following statement:

The closure, in $K^S$, of the set $\text{conv } W$ contains a function $w_0$ such that $w_0(s) \geq \epsilon$ for all $s \in S$. 
We have seen that the existence of such a function implies the inequality $e(W) \geq e$. The lemma may thus be reformulated in the following form

$$e(W) = \sup_{b \in P(W)^-} \inf_{s \in S} b(s).$$

It is the purpose of the present note to prove that we have, in fact, equality. The proof of the other inequality seems to require the compactness of $K^S$. We state it in the form of a minimax theorem for a relation on $W \times S$; in other words, for a function assuming only the values 1 and 0 on $W \times S$.

**Theorem.** Given any relation on $W \times S$, in other words any family $W$ of subsets of $S$, we have

$$\inf_{\lambda \in P(S)} \sup_{w \in W} \lambda(w) = \sup_{w \in P(W)^-} \inf_{s \in S} w(s).$$

**Proof.** It remains to prove the inequality

$$e(W) \leq \sup_{w \in P(W)^-} \inf_{s \in S} w(s).$$

Let us show that, for every finite $F \subset S$ there exists a $b \in \text{conv } W$ such that $b(s) \geq e(W)$ for all $s \in F$. Define a mapping $G$ of $P(S)$ into $R^n$ by the formula

$$G(x) = (x(s_1), \ldots, x(s_n)).$$

Denote by $M$ the subset of $R^n$ consisting of all $(y_1, \ldots, y_n) \in R^n$ for which all $y_i \geq e(W)$. Suppose the intersection $G(\text{conv } W) \cap M$ is void. Since $G(\text{conv } W)$ is compact there exists a linear form $\alpha$ on $R^n$,

$$\alpha(x) = \alpha_1 x_1 + \cdots + \alpha_n x_n$$

such that $\sup \alpha(G(\text{conv } W)) < \inf \alpha(M)$. Since $\inf \alpha(M)$ is finite, it follows that all $\alpha_i$ are nonnegative. Since $\alpha$ is nonzero we may assume that $\sum \alpha_i = 1$, in other words $\alpha \in P(S)$. Now $\inf \alpha M \leq e(W)$ so that

$$\sup \alpha(W) = \sup \alpha(\text{conv } W) = \sup \alpha(G \text{ conv } W) < \inf \alpha M \leq e(W),$$

a contradiction. The proof is complete. □

3. **Concluding remarks**

The combinatorial lemma was published in 1959 and was the result of an effort to isolate the combinatorial substance of the theory of weak compactness, in particular to explain why conditions of countable character already imply compactness and to eliminate complicated measure theory from the proof that weak compactness extends from a set to its convex hull. As one of the first applications, the author gave a proof of the following proposition: *If a subset $M$...*
of a Banach space satisfies the double limit condition, then so does $\text{conv } M$. This proof is reproduced in the monograph of Köthe [4, 5]. A systematic account of the applications of the lemma is to be found in the author's lecture at the 1961 convexity symposium in Seattle [8]. The most convenient form in which the lemma can be applied is a Fubini-like result which can also be stated as an extension theorem for separately continuous functions. This theorem [7, 9] contains the Eberlein theorem as well as the theorem on convex hulls.

At the 1965 convexity symposium in Copenhagen the author suggested [10] the possibility of applying these ideas to game theory. A number of subsequent papers, notably by Young [13, 14] and Kindler [2, 3] contained further contributions towards clarifying the connections with game theory, in particular the connections with the minimax theorem. The double limit condition was also the subject of further investigations of Simons [11, 12].

The combinatorial lemma gives conditions for $\inf \sup \lambda(w) = 0$.

For applications in analysis we are interested in families $W$ such that this relation is satisfied not only for the family $W$ itself but also for all families of the form $W \cap R$, $R$ being an arbitrary infinite subset of $S$. We denote by $W \cap R$ the family of all sets $w \cap R$, $w \in W$. In other words, we postulate for each infinite $R \subset S$, the equality

$$\inf \sup (w) = 0$$

as $\lambda$ ranges over all $\lambda \in P(S)$ with carrier contained in $R$. This corresponds to the fact that the lemma is to be applied to questions concerning convergence of sequences and these questions, in their turn, may be reformulated in terms of the behaviour of all subsequences.

Of course, the quantity $\inf \sup \lambda(w)$ is meaningful for an arbitrary family $W$. This quantity, called the thickness of $W$ in [1], represents a useful numerical characteristic of the combinatorial structure of the family. In spite of the fact that it does not seem to have any immediate use for applications in functional analysis it nevertheless deserves to be considered in its own right. It has an interesting geometrical interpretation; we have shown, in [1], that every number between zero and one may be represented as the thickness of a suitable family $W$. The present note was motivated by an attempt to give a characterization of families of thickness zero, an attempt that resulted in obtaining the dual description of the notion of thickness. The method based on duality makes it possible to consider, together with the family $W$, also the family of functions $\text{conv } W$. This is best illustrated by the juxtaposition of the conditions for families $W$ of thickness zero and those for which $e(W) = 0$ in the stronger sense. More precisely we shall consider families $W$ with the following properties
(1) \( \varepsilon(W) = 0 \) hereditarily, in other words: for each infinite \( R \subset S \)
\[
\inf_{\lambda} \sup_{w} (w) = 0
\]
as \( \lambda \) ranges over all \( \lambda \in P(S) \) with carrier in \( R \).

(2) \( \varepsilon(W) = 0 \).

The characterizations seem to be more accessible in their negative form; they are as follows.

(1') The closure of \( W \) contains a function \( b \) such that \( b(s) = 1 \) for all \( s \in R \), an infinite subset of \( S \).

(2') The closure \( \text{conv} W \) contains a function \( b \) such that \( b(s) \geq \varepsilon \) for all \( s \in S \) and a positive number \( \varepsilon \).

References