SIAM J. OPTIMIZATION
Vol. 6, No. 4, pp. 1153–1162, November 1996

015

OPTIMALITY OF NESTED PARTITIONS AND ITS APPLICATION TO CLUSTER ANALYSIS*

E. BOROS† AND F. K. HWANG‡

Abstract. A partition of a set N of n distinct numbers is called nested if four numbers $a < b < c < d$ in N such that a and c are in one part while b and d in another do not exist. A partition is called a p-partition if the number of parts is specified at p and a shape-partition if the sizes of the p parts are also specified. There are exponentially many p-partitions but only polynomially many nested p-partitions. In this paper we consider these notions in d-dimensional Euclidean spaces and give a general condition on the cost structure for which an optimal shape-partition is always nested. We illustrate applications of our results to some clustering problems, generalize some known results in this way, and propose some open problems.

Key words. clustering, nested partitions

AMS subject classifications. 62H30, 05A18

1. Introduction.

Consider the problem of partitioning a set N of n distinct numbers into nonempty disjoint parts. The partition is called an *open-partition* if the number of parts is not prespecified and called a *p-partition* if the number is specified to be p. If, furthermore, a set $\{n_1, \ldots, n_p\}$ with $\sum_{i=1}^{p} n_i = n$ is prespecified to be the set of sizes of the p parts, then the partition is called a *shape-partition*, shape referring to the set $\{n_i\}$.

Often, one encounters the problem of finding an optimal partition given a cost (of partition) function. However, the brute force approach of comparing the costs of all partitions is too time-consuming due to the large number of partitions. For example, using the principle of inclusion–exclusion, the number of p-partitions can be shown to be

$$\#(n,p) = \frac{1}{p!}\sum_{k=0}^{p-1}(-1)^k\binom{p}{k}(p-k)^n. \tag{1}$$

The number of open-partitions

$$\#(n) = \sum_{p=1}^{n}\#(n,p) \tag{2}$$

is represented by the *Bell numbers* whose first 10 terms are 1, 2, 5, 15, 52, 203, 877, 4140, 21147, 115975. Even for the shape-partition, the number is

$$\#(n_1,\ldots,n_p) = \frac{n!}{\prod_{i=1}^{p} n_i! \prod_{j=1}^{n-p+1} p_j!}, \tag{3}$$

where p_j is the number of parts of size j. This number is easily seen to be exponential in n even for $p = 2$.

* Received by the editors August 4, 1994; accepted for publication (in revised form) April 20, 1995. This paper is a revised version of RUTCOR Research Report #7, 1993.

† RUTCOR, Rutgers University, New Brunswick, NJ 08904 (boros@rutcor.rutgers.edu). The research of this author was supported in part by Office of Naval Research grants N00014-92-J-1375 and N00014-92-J-4083 and by Air Force Office of Scientific Research grant F49620-95-1-0233.

‡ Department of Applied Mathematics, Chiao-Tung University, Hsin-Chu, Taiwan 30050 ROC (fhwang@math.nctu.edu.tw).

One way to deal with the combinatorial problem of huge partition spaces is to look for small subspaces which, nevertheless, also contain optimal partitions. One well-studied subspace consists of *consecutive partitions* [3,9] which are characterized by the requirement that each part of a consecutive partition consists of numbers consecutive in N. In this subspace each p-partition corresponds to a way of inserting $p-1$ bars into the $n-1$ spaces between the n numbers. The number of p-partitions is thus

$$\#_C(n,p) = \binom{n-1}{p-1}, \tag{4}$$

a polynomial function of n for fixed p. For shape-partitions, the number is easily seen to be

$$\#_C(n_1,\ldots,n_p) = \frac{p!}{\prod_{j=1}^{n-p+1} p_j!}. \tag{5}$$

When the "consecutive" subspace is not known to contain an optimal partition, one has to search other subspaces. Boros and Hammer [2] raised the notion of *nested partitions*, which is defined by the nonexistence of four numbers $a<b<c<d$ in N such that a and c belong to one part, while b and d belong to another. Note that a consecutive partition is always nested, but not vice versa. They demonstrated some cost functions which guarantee that an optimal p-partition is nested. Hwang and Mallows [10] showed that the number of nested p-partitions is

$$\#_N(n,p) = \frac{\binom{n}{p-1}\binom{n}{p}}{n}, \tag{6}$$

again, a polynomial function of n for fixed p.

The notions of "consecutiveness" and "nestedness" have been extended to vectors (points in d-dimensional spaces). We now extend them further to d-dimensional multisets. Let $X=\{x_1,x_2,\ldots,x_n\}\subset\mathbb{R}^d$ be a *multiset* of d-dimensional points, i.e., elements of X may coincide. Furthermore, let $\mathrm{conv}(X)$ denote the *convex hull* of X, and let $\mathrm{conv}^*(X)$ denote the relative interior of $\mathrm{conv}(X)$. A partition $\pi=(\pi_1,\ldots,\pi_p)$ of the multiset X (identical points are treated as separated entities in a partition) is called *consecutive* (see [1]) if $\mathrm{conv}^*(\pi_i)\cap\mathrm{conv}^*(\pi_j)=\emptyset$ for all $1\le i,j\le p$. It is called *nested* (see [2]) if for all $1\le i,j\le p$, either $\pi_i\cap\mathrm{conv}^*(\pi_j)=\emptyset$ or $\pi_j\cap\mathrm{conv}^*(\pi_i)=\emptyset$. Again, consecutiveness implies nestedness. In this paper we give a sufficient condition on the cost function such that an optimal shape-partition is always nested. A by-product is a sufficient condition for the existence of a consecutive optimal shape-partition. In particular, they lead to an extension of Fisher's result [6] on a clustering problem from one dimension to d dimensions which he long desired.

2. The main results. In this section we derive a general condition which guarantees that every optimal shape-partition is nested. We first consider a shape with only two parts, and then we extend the result to general p.

Consider a multiset X of d-dimensional points, $d\ge 1$, and a partition of it into two parts, $\pi=(\pi_1,\pi_2)$. Let, furthermore, $F(\pi)$ denote the cost of partition π.

Let π' be a partition obtained from π by interchanging two points, $x\in\pi_1$ and $y\in\pi_2$. Clearly, π' has the same shape as π. We will consider $F(\pi')-F(\pi)$ as a function $\Delta_F(x,y)$ of x and y, i.e., $\Delta_F:\pi_1\times\pi_2\longrightarrow\mathbb{R}$. More precisely, let us consider a continuous, real-valued mapping Δ_F^* over the space $\mathbb{R}^d\times\mathbb{R}^d$ satisfying the following

conditions:

$$\begin{array}{ll} \Delta_F^*(x,y) = \Delta_F(x,y) & \text{for all } x \in \pi_1 \text{ and } y \in \pi_2, \\ \Delta_F^*(z,z) = 0 & \text{for all } z \in \mathbb{R}^d. \end{array} \tag{7}$$

Such a mapping Δ_F^* exists and can naturally be considered as a continuous extension of Δ_F, for if $x \in \pi_1$ and $y \in \pi_2$ happen to coincide (X is a multiset), then $\Delta_F(x,y) = 0$ since the switch of identical elements does not change the partition. Let us remark that in most cases, when F is given in an algebraic form, the formula for Δ_F will automatically define such an extension.

For a fixed vector $x \in \pi_1$ let us introduce the notation $g_x^*(y) = \Delta_F^*(x,y)$ and, analogously, let $g_y^*(x) = \Delta_F^*(x,y)$, if we want to emphasize that $y \in \pi_2$ is fixed now. Let, furthermore, $X^+(g_y^*) = \{x \in \mathbb{R}^d | g_y^*(x) \geq 0\}$, and let $Y^+(g_x^*) = \{y \in \mathbb{R}^d | g_x^*(y) \geq 0\}$.

We are ready now to state a sufficient condition for a shape-partition to be nested.

THEOREM 2.1. *For a shape-partition problem let $\pi = (\pi_1, \pi_2)$ be an optimal partition. Further let us suppose that either for every $x \in \pi_1$, the set $Y^+(g_x^*)$ is a convex set with x being a boundary point, or for every $y \in \pi_2$, the set $X^+(g_y^*)$ is a convex set with y being a boundary point. Then $\pi = (\pi_1, \pi_2)$ is nested.*

Proof. Let us assume that for every $y \in \pi_2$, the set $X^+(g_y^*)$ is a convex set with y being a boundary point. Since π is optimal, i.e., its cost $F(\pi)$ is minimal among all partitions of the same shape, $g_y^*(x) \geq 0$ for all $x \in \pi_1$ and all $y \in \pi_2$, implying

$$\pi_1 \subset X^+(g_y^*) \tag{8}$$

for all $y \in \pi_2$. Since for each $y \in \pi_2$ the set $X^+(g_y^*)$ is convex with y being a boundary point, the intersection of all these sets

$$X^+ = \bigcap_{y \in \pi_2} X^+(g_y^*)$$

is also convex and no point of π_2 belongs to its interior. Since $\pi_1 \subset X^+$ by (8), $\pi_2 \cap \text{conv}^*(\pi_1) = \emptyset$ follows, which proves that π is nested. The other case is analogous. □

Sometimes, it is easier to use Theorem 2.1 when the conditions are specified on the functions g_y^* and g_x^*.

A real-valued function $f(x)$ is called *quasi concave* if over any interval $[a,b] = \{\alpha a + (1-\alpha)b | 0 \leq \alpha \leq 1\}$ it always attains its minimum over $[a,b]$ at one of the endpoints. The function f is called *strictly quasi concave* if no internal point of an interval can be a minimum (over that interval). It is well known that a (strictly) concave function is (strictly) quasi concave.

COROLLARY 2.2. *Let X be a given multiset with F being the cost function of its partitions, as before. If either g_x^* for any $x \in X$ or g_y^* for any $y \in X$ is strictly quasi concave, then every optimal shape-partition is nested.*

Proof. Let us assume that g_y^* for any $y \in X$ is strictly quasi concave. The other case can be treated analogously.

Let us consider an optimal shape-partition $\pi = (\pi_1, \pi_2)$. According to the previous theorem, if $X^+(g_y^*)$ is convex having y on its boundary for every $y \in \pi_2$, then π is necessarily nested.

Let us observe first that for every $y \in \pi_2$, the point y must be a boundary point of $X^+(g_y^*)$, since g_y^* is strictly quasi concave. This implies that if π is not nested,

then, by Theorem 2.1, there is a vector $y \in \pi_2$ for which the set $X^+(g_y^*)$ is not convex. Then there must exist points $u, v \in X^+(g_y^*)$ and $w = \alpha u + (1-\alpha)v \notin X^+(g_y^*)$ for some $0 < \alpha < 1$, i.e., for which $g_y^*(u) \geq 0$, $g_y^*(v) \geq 0$ while $g_y^*(w) < 0$. Since g_y^* is continuous, the interval $[u, v]$ has an internal minimum, contradicting the strict quasi concavity of g_y^*. □

THEOREM 2.3. *Suppose that the cost function has the structure* $F(\pi) = \sum_{i=1}^{p} f(\pi_i)$, *i.e.,* $F(\pi)$ *is the sum of independent values associated with each of the parts. In this case, if every optimal shape-partition is nested holds for* $p = 2$, *then it holds for* $p > 2$.

Proof. Let π be an optimal shape-partition. By Theorem 2.1, any two parts of π must be a nested partition of their elements or we would be able to reduce $F(\pi)$ by making them nested, which contradicts the assumption that π is optimal. By the definition of a nested partition, π is nested if any two parts of π are pairwise nested. □

Interestingly, the arguments used to establish nested optimal partitions are also applicable for consecutive optimal partitions, for which more efficient algorithms exist.

THEOREM 2.4. *Consider a shape-partition problem and let* π *be an optimal partition. Suppose that for every pair* (π_i, π_j) *and for every* $x \in \pi_j$, $Y^+(g_x^*)$ *is a convex set with* x *on its boundary, and for every* $y \in \pi_j$, $X^+(g_y^*)$ *is a convex set having* y *as a boundary point. Then every optimal partition is consecutive.*

Proof. First consider the case of two parts. Let π be an optimal partition. We have argued in the proof of Theorem 2.1 that $X^+(g_y^*)$ being a convex set having y on its boundary implies that no $y \in \pi_2$ is in $\text{conv}^*(\pi_1)$. Similarly, $Y^+(g_x^*)$ being a convex set with x being the boundary implies that no $x \in \pi_1$ is in $\text{conv}^*(\pi_2)$. Hence π is consecutive. The result is then extended to general p parts by an argument analogous to the proof of Theorem 2.3. □

COROLLARY 2.5. *Suppose that both* g_x^* *and* g_y^* *are strictly quasi concave for every* x *and* y, *respectively. Then every optimal shape-partition is consecutive.*

Since an open-partition must be a p-partition for some p, and a p-partition must be a shape-partition for some shape, results in this section also apply to p-partitions and open-partitions.

3. Applications to clustering. In a clustering problem, one partitions a given set of points into clusters usually with points in the same cluster close to each other, though closeness can be defined in various ways. It is very rare for a clustering problem to have a polynomial-time algorithm for exact optimal clustering, due to the usually large number of possible clusterings. One of the few exceptions is due to Fisher who was one of the first to use consecutive partitions. Fisher [6] considered a one-dimensional clustering problem where the goal is to minimize the sum of squares, i.e., the cost of a partition $\pi = (\pi_1, \ldots, \pi_p)$ is

$$F(\pi) = \sum_{i=1}^{p} \sum_{x_j \in \pi_i} (x_j - \overline{x}_i)^2, \tag{9}$$

where $\overline{x}_i$ is the average of the numbers in π_i. He proved that there exists a consecutive optimal p-partition, even when there is a weight w_j associated to each number x_j. Since every open-partition must be a p-partition for some p, this also implies the existence of a consecutive optimal open-partition. Fisher wrote [6, pp. 796–797]: "It would of course be most desirable to develop, both theoretically and computationally, a distance criterion that is defined in more than one dimension. An example of the need for such a formulation is shown in a multivariate stratification problem

encountered in a sample survey by Hagood and Bernert [8]. Of course involved in any such approach is a relevant system of weighing the different dimensions to reflect their relative importance in determining distance." Gower [7] studied three criteria commonly adopted in the literature of cluster analysis for multivariate data. One of which, attributed to Edwards and Cavalli-Sforza [5], is to divide the data into two disjoint subsets with a minimum sum of squares, a special case of Fisher's d-dimensional problem with $p = 2$.

Unfortunately, Fisher's proof technique of the one-dimensional case cannot handle a weight function associated with the dimensions. Gower proved the existence of consecutive optimal partitions for $p = 2$ and without dimension weight. We now consider the general case. Suppose that $x_j = (x_{j1}, \ldots, x_{jd})$ and u_k is the positive weight of dimension k, $k = 1, \ldots, d$. Consider the cost function

$$F(\pi) = \sum_{i=1}^{p} \sum_{x_j \in \pi_i} \sum_{k=1}^{d} u_k (x_{jk} - \overline{x}_{ik})^2 = \sum_{i=1}^{p} \sum_{x_j \in \pi_i} (y_j - \overline{y}_i)^2, \tag{10}$$

where $y_j = (\sqrt{u_1} x_{j1}, \sqrt{u_2} x_{j2}, \ldots, \sqrt{u_d} x_{jd})$ and $\overline{y}_i = \sum_{x_j \in \pi_i} y_j / |\pi_i|$ is the mean (centroid) of the vectors y_j for $x_j \in \pi_i$ (and where the product of the vectors denotes their inner product).

In the following theorem we shall replace y_j in (10) by x_j for uniformity and also generalize it by introducing a weight function w_i associated to part i.

THEOREM 3.1. *Suppose that*

$$F(\pi) = \sum_{i=1}^{p} w_i \sum_{x_j \in \pi_i} (x_j - \overline{x}_i)^2, \tag{11}$$

where $w_i > 0$ and $\overline{x}_i$ is the centroid (mean, in this case) of the d-dimensional points in π_i. Then an optimal shape-partition must be nested.

Proof. By Theorem 2.3 it is enough to prove the above statement for the case of $p = 2$.

Let π be an optimal shape-partition and let π' be the partition obtained from π by interchanging $y \in \pi_1$ and $z \in \pi_2$. Let $\overline{x}'_1$ and $\overline{x}'_2$ denote the centroids of π'_1 and π'_2, respectively. Then

$$0 \le \Delta_F(y, z) = w_1 \left[\sum_{x_j \in \pi'_1} (x_j - \overline{x}'_1)^2 - \sum_{x_j \in \pi_1} (x_j - \overline{x}_1)^2 \right]$$

$$+ w_2 \left[\sum_{x_j \in \pi'_2} (x_j - \overline{x}'_2)^2 - \sum_{x_j \in \pi_2} (x_j - \overline{x}_2)^2 \right]$$

$$= w_1 \left[\sum_{x_j \in \pi'_1} x_j^2 - \frac{\left(\sum_{x_j \in \pi'_1} x_j \right)^2}{n_1} - \sum_{x_j \in \pi_1} x_j^2 + \frac{\left(\sum_{x_j \in \pi_1} x_j \right)^2}{n_1} \right]$$

$$
\begin{aligned}
&+ w_2 \left[\sum_{x_j \in \pi_2'} x_j^2 - \frac{\left(\sum_{x_j \in \pi_2'} x_j\right)^2}{n_2} - \sum_{x_j \in \pi_2} x_j^2 - \frac{\left(\sum_{x_j \in \pi_2} x_j\right)^2}{n_2} \right] \\
&= w_1 \left[z^2 - y^2 - \frac{(z-y)^2 + 2(z-y)\sum_{x_j \in \pi_1} x_j}{n_1} \right] \\
&\quad + w_2 \left[y^2 - z^2 - \frac{(y-z)^2 + 2(y-z)\sum_{x_j \in \pi_2} x_j}{n_2} \right] \\
&= w_1 \left[z^2 - y^2 - \frac{(z-y)^2}{n_1} - 2(z-y)\overline{x}_1 \right] + w_2 \left[y^2 - z^2 - \frac{(y-z)^2}{n_2} - 2(y-z)\overline{x}_2 \right] \\
&= \left(w_1 - w_2 - \frac{w_1}{n_1} - \frac{w_2}{n_2} \right) z^2 + \left(w_2 - w_1 - \frac{w_1}{n_1} - \frac{w_2}{n_2} \right) y^2 \\
&\quad + 2 \left(\frac{w_1}{n_1} + \frac{w_2}{n_2} \right) yz + 2(w_2 \overline{x}_2 - w_1 \overline{x}_1) z + 2(w_1 \overline{x}_1 - w_2 \overline{x}_2) y.
\end{aligned}
$$

View the above expression as a function of real y and z (vectors) with the given coefficient ($\overline{x}_1$ and $\overline{x}_2$ are treated as fixed) and define $g_z^*(y)$ and $g_y^*(z)$ accordingly. Since the sum of the coefficients of the z^2 term and the y^2 term is negative, at least one of them is negative, say, the coefficient of the z^2 term. Since $g_z^*(y)$ is separable in the dimension of y, it is easily verified that the negative coefficient of the z^2 term implies that the Hessian is negative-definite. Hence g_z^* is strictly concave. Since the coefficients of y^2 and z^2 are independent of the particular selection of y and z, we can conclude that g_z^* is strictly concave for all $z \in \pi_2$. By Corollary 2.2 an optimal shape-partition thus must be nested. □

COROLLARY 3.2. *If*

$$
|w_1 - w_2| \leq \frac{w_1}{n_1} + \frac{w_2}{n_2}, \tag{12}
$$

then every optimal shape-partition is consecutive.

Proof. The proof of Corollary 3.2 follows immediately from Corollary 2.5. □

In particular, if $w_i = 1$ for all i, then the condition of Corollary 3.2 is satisfied. Thus we have extended Fisher's sum-of-squares result to d-dimensional points.

By setting $w_i = 0$ for $n_i = 1$ and $w_i = 1/(n_i - 1)$ for $n_i \geq 2$, $F(\pi)$ in Theorem 3.1 represents the sum of variances (for multidimensional points, each variance is a weighted sum over the d dimensions). Therefore all shape-partitions to minimize the sum of variances are nested. It is also easily verified that (12) holds if $|n_1 - n_2| \leq 1$. Hence we have the following result.

COROLLARY 3.3. *Consider a partitioning problem where the part-sizes can differ by at most* 1. *Then every optimal partition minimizing the sum of variances is consecutive.*

One may feel that perhaps for an arbitrary shape there exists a consecutive optimal partition. We now give a one-dimensional example to show that Corollary 3.3 is tight, i.e., if the part-size can differ by 2, then no optimal partition is consecutive.

Let $N = \{0, 13, 14, 14, 15, 28\}$ and the shape be $\{2, 4\}$. Then $\pi_1 = \{14, 14\}$, $\pi_2 = \{0, 13, 15, 28\}$ is the optimal shape-partition minimizing the sum of variances. But $\{\pi_1, \pi_2\}$ is not a consecutive partition.

Another consequence of Corollary 2.2 is a strengthening of Theorem 1.2 of [2].

THEOREM 3.4. *Suppose that*

$$F(\pi) = \sum_{i=1}^{p} w_i \sum_{x_j, x_k \in \pi_i} (x_j - x_k)^2, \tag{13}$$

where $w_i > 0$. *Then every optimal shape-partition is nested.*

Proof. By Theorem 2 again, it is enough to consider $p = 2$. Let $y \in \pi_1$ and $z \in \pi_2$. Then

$$\begin{aligned} \Delta_F(y,z) &= w_1 \sum_{x_j \in \pi_1} \left((z - x_j)^2 - (y - x_j)^2\right) \\ &\quad + w_2 \sum_{x_j \in \pi_2} \left((y - x_j)^2 - (z - x_j)^2\right) \\ &= y^2[w_2(n_2 - 1) - w_1(n_1 + 1)] + z^2[w_1(n_1 - 1) - w_2(n_2 + 1)] \\ &\quad + 2(w_1 + w_2)yz + 2y\left[w_1 \sum_{x_j \in \pi_1} x_j - w_2 \sum_{x_j \in \pi_2} x_j\right] \\ &\quad + 2z\left[w_2 \sum_{x_j \in \pi_2} x_j - w_1 \sum_{x_j \in \pi_1} x_j\right], \end{aligned} \tag{14}$$

where n_1 and n_2 denote the cardinalities of π_1 and π_2, respectively. Since the sum of the coefficients of y^2 and z^2 is $-2(w_1 + w_2) < 0$, at least one of them is negative, implying that at least one of g_z^* or g_y^* is strictly concave. Thus, by Corollary 2.2, we can conclude that an optimal shape-partition must be nested. □

If both y^2 and z^2 have nonpositive coefficients in the above proof, i.e., if

$$|w_1 n_1 - w_2 n_2| \le w_1 + w_2, \tag{15}$$

then by Corollary 2.5 an optimal shape-partition must be consecutive. This observation yields the following interesting consequence.

COROLLARY 3.5. *Consider a shape-partitioning problem where*

$$F(\pi) = \sum_{i=1}^{p} \sum_{x_j, x_k \in \pi_i} (x_j - x_k)^2 \tag{16}$$

and in which the part-sizes can differ by at most 2. *Then every optimal shape-partition is consecutive.*

Boros and Hammer studied a one-dimensional clustering problem with

$$F(\pi) = \sum_{i=1}^{p} \sum_{x_j, x_k \in \pi_i} |x_j - x_k| \tag{17}$$

and proved that every optimal p-partition is nested. This, however, may not be true for shape-partitions. In this paper, instead of the absolute difference between two numbers, we consider the absolute difference between a number and the centroid of the part, with a part-weight.

THEOREM 3.6. *Suppose that*

$$F(\pi) = \sum_{i=1}^{p} w_i \sum_{x_j \in \pi_i} |x_j - m_i|, \tag{18}$$

$$z...z \;\; y...y \;\; |_{m_1} \;\; y...y \;\; z...z \;\; |_{m_2} \;\; y...y \; z...z$$

FIG. 1. *The ordering of y's and z's when* $w_1 \geq w_2$.

where $w_i > 0$ *and* m_i *is the median of the set* π_i. *Then every optimal shape-partition is nested.*

Proof. By Theorem 2.3 it is sufficient to consider $p = 2$. Let $y \in \pi_1$ and $z \in \pi_2$. Let $\pi' = (\pi'_1, \pi'_2)$ be obtained from π by interchanging y and z, and let m'_1 and m'_2 be the medians of π'_1 and π'_2. Without loss of generality, assume $m_1 \leq m_2$.

Case i. $y, z \geq m_2$. Then $m'_1 = m_1$ and $m'_2 = m_2$.

$$\begin{aligned} 0 \leq \Delta_F(y,z) &= w_1(z - m_1) + w_2(y - m_2) - w_1(y - m_1) - w_2(z - m_2) \\ &= (w_1 - w_2)(z - y). \end{aligned} \tag{19}$$

So $z - y$ has the same sign as $w_1 - w_2$.

Case ii. $m_1 \leq y, z \leq m_2$. Then $m'_1 = m_1$ and $m'_2 = m_2$.

$$\begin{aligned} 0 \leq \Delta_F(y,z) &= w_1(z - m_1) + w_2(m_2 - y) - w_1(y - m_1) - w_2(m_2 - z) \\ &= (w_1 + w_2)(z - y). \end{aligned} \tag{20}$$

So $z \geq y$.

Case iii. $y, z \leq m_1$. Then $m'_1 = m_1$ and $m'_2 = m_2$.

$$\begin{aligned} 0 \leq \Delta_F(y,z) &= w_1(m_1 - z) + w_2(m_2 - y) - w_1(m_1 - y) - w_2(m_2 - z) \\ &= (w_1 - w_2)(y - z). \end{aligned} \tag{21}$$

So $y - z$ has the same sign as $w_1 - w_2$.

First consider $w_1 \geq w_2$, then the ordering of $y \in \pi_1$ and $z \in \pi_2$ in the three intervals separated by m_1 and m_2 is shown in Figure 1.

We will show that a $y \in \pi_1$, $y \geq m_2$ and a $z \in \pi_2$, $m_1 \leq z < m_2$ cannot coexist. This will imply that the partition is nested. Suppose to the contrary that such a pair (y, z) exists. Then

$$\begin{aligned} 0 \leq \Delta_F(y,z) &= w_1 \sum_{x_j \in \pi'_1} |x_j - m'_1| + w_2 \sum_{x_j \in \pi'_2} |x_j - m'_2| \\ &\quad - w_1 \sum_{x_j \in \pi_1} |x_j - m_1| - w_2 \sum_{x_j \in \pi_2} |x_j - m_2| \\ &\leq w_1 \sum_{x_j \in \pi'_1} |x_j - m_1| + w_2 \sum_{x_j \in \pi'_2} |x_j - m_2| \\ &\quad - w_1 \sum_{x_j \in \pi_1} |x_j - m_1| - w_2 \sum_{x_j \in \pi_2} |x_j - m_2| \\ &= w_1(z - m_1) + w_2(y - m_2) - w_1(y - m_1) - w_2(m_2 - z) \\ &= w_1(z - y) + w_2(y + z - 2m_2) \\ &< w_2(2z - 2m_2) < 0, \end{aligned}$$

an absurdity.

Next consider $w_1 \leq w_2$; then the possible ordering of the y's and z's is as shown in Figure 2.

$$y...y \quad z...z \quad |_{m_1} \quad y...y \quad z...z \quad |_{m_2} \quad z...z \quad y...y$$

FIG. 2. *The ordering of y's and z's when* $w_1 \leq w_2$.

We can show in a similar fashion that $z \leq m_1$ and $m_1 < y \leq m_2$ cannot coexist. Thus the partition is nested. □

One can also observe that if $w_1 = w_2$, then the only order of the elements of π_1 and π_2 satisfying all conditions in the above proof is $y \mid y \;\; z \mid z$, implying hence the following corollary (for a different proof, see also [9]).

COROLLARY 3.7. *If*

$$F(\pi) = \sum_{i=1}^{p} \sum_{x_j \in \pi_i} |x_j - m_i|, \tag{22}$$

where m_i *denotes the median of the set* π_i, *then every optimal shape-partition is consecutive.*

4. Some concluding remarks. Boros and Hammer proved that if $F(\pi) = \sum_{i=1}^{p} \sum_{x,y \in \pi_i} |x - y|$ for one-dimensional points or if $F(\pi) = \sum_{i=1}^{p} \sum_{x,y \in \pi_i} (x - y)^2$ for d-dimensional points, then an optimal partition is nested. We generalized their result by giving a broad sufficient condition derived from a novel geometric argument. We applied this condition to obtain the d-dimensional version of Fisher's clustering problem which he proposed but couldn't prove.

Since the concept of nested partition is fairly recent, there are still many unexplored issues. We raise the following questions.

(i) Clearly, a partition $\pi = (\pi_1, \ldots, \pi_p)$ is nested if and only if for any i and j, (π_i, π_j) is a nested partition of the elements in $\pi_i \cup \pi_j$. Does the existence of a nested optimal 2-partition guarantee the existence of a nested optimal p-partition for general p? An affirmative answer was recently given by Hwang, Rothblum, and Yao [11] for one-dimensional points but the problem for general dimension remains open.

(ii) For the subspace of consecutive partitions (in one dimension), there exist an $O(n^2)$-time dynamic programming algorithm to find an optimal open-partition [9] and an $O(pn^2)$-time algorithm to find an optimal p-partition. From (6), there exists an $O(n^{2p-2})$-time algorithm to find an optimal nested p-partition. Does there exist a better dynamic programming algorithm for the subspace of nested p-partitions? Note that since any subset can be a part in a nested open partition, for a general cost function, one must inspect at least 2^n cost terms to find an optimal nested open-partition.

For the one-dimensional case, it seems quite plausible to conjecture that there always exists a consecutive optimal partition. However, a counterexample was recently given by Chang and Hwang [4] in which the optimal partition is nested but is not consecutive.

REFERENCES

[1] E. R. Barnes, A. J. Hoffman, and U. G. Rothblum, *On optimal partitions having disjoint convex and conic hulls*, Math. Programming, 54 (1992), pp. 69–86.

[2] E. Boros and P. L. Hammer, *On clustering problems with connected optima in Euclidean spaces*, Discrete Math., 75 (1989), pp. 81–88.

[3] A. K. Chakravarty, J. B. Orlin, and U. G. Rothblum, *A partitioning problem with additive objective with an application to optimal inventory grouping for joint replenishment*, Oper. Res., 30 (1982), pp. 1018–1022.

[4] G. J. Chang and F. K. Hwang, *Optimality of consecutive and nested tree partitions*, to appear.

[5] A. W. F. Edwards and L. L. Cavalli-Sforza, *A method for cluster analysis*, Biometrics, 21 (1965), pp. 362–375.

[6] W. D. Fisher, *On grouping for maximum homogeneity*, J. Amer. Statist. Assoc., 53 (1958), pp. 789–798.

[7] J. C. Gower, *Some distance properties of latent roots and vector methods used in multivariate analysis*, Biometrika, 53 (1966), pp. 325–338.

[8] M. J. Hagood and E. H. Bernert, *Component indexes as a basis for stratification in sampling*, J. Amer. Statist. Assoc., 40 (1945), pp. 330–341.

[9] F. K. Hwang, *Optimal partitions*, J. Optim. Theory Appl., 34 (1981), pp. 1–10.

[10] F. K. Hwang and C. L. Mallows, *The numbers of nested partitions and inner-consecutive partitions*, J. Combin. Theory Ser. A, 70 (1995), pp. 323–333.

[11] F. K. Hwang, U. G. Rothblum, and Y. C. Yao, *Localizing combinatorial properties of partitions*, Discrete Math., to appear.