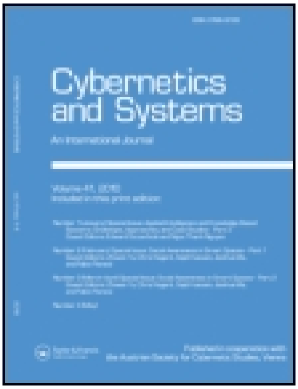

Journal of Cybernetics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ucbs19

Well-Separated Clusters and Optimal Fuzzy Partitions

J. C. Dunn†
[a] Cornell University Department of Theoretical and Applied Mechanics, and The Center for Applied Mathematics
Published online: 30 Apr 2008.

Journal of Cybernetics, 1974, 4, 1, 95-104

Well-Separated Clusters and Optimal Fuzzy Partitions

J. C. Dunn†
Cornell University
Department of Theoretical and Applied Mechanics,
and The Center for Applied Mathematics

Abstract

Two separation indices are considered for partitions $P = \{X_1, \ldots, X_k\}$ of a finite data set X in a general inner product space. Both indices increase as the pairwise distances between the subsets X_i become large compared to the diameters of X_i. Maximally separated partitions P' are defined and it is shown that as the indices of P' increase without bound, the characteristic functions of X_i' in P' are approximated more and more closely by the membership functions in fuzzy partitions which minimize certain fuzzy extensions of the k-means squared error criterion function.

1. Introduction

In reference [1], the problem of detecting the presence or absence of compact well separated (CWS) clusters in a large finite data set X was approached by considering a related extremization problem for fuzzy extensions of the standard k-means squared error criterion function. When k CWS clusters $\{X_1, \ldots, X_k\}$ are present in X it is plausible that the membership functions in a globally minimizing fuzzy partition should approximate the characteristic functions of the clusters X_i more and more closely as the separation between clusters increases. On the other hand, when well-separated clusters are not present in X, one expects the minimizing fuzzy partition to consist of membership functions which are truly fuzzy in the sense that their values differ substantially from 0 or 1 over certain regions in X. These conjectures were supported by a number of specific examples described in [1], however no formal proofs were available at that time. In this paper, we formulate a precise statement and rigorous proof of the first conjecture.

2. Separation Indices

Let X denote a finite subset of a general real inner product space V, i.e., a real vector space equipped with a symmetric positive definite bi-linear real function, $\langle u|v\rangle$ defined for all pairs of vectors u, v in V [2]. The inner product $\langle u|v\rangle$ induces a metric d on V according to

†Requests for reprints should be sent to J.C. Dunn, Department of Theoretical and Applied Mechanics, Cornell University, College of Engineering, Ithaca, New York 14850.

the rule,

$$d(u, v) \triangleq \|u - v\| \triangleq \langle u - v | u - v\rangle^{1/2}$$

With respect to this distance we may define diameters of subsets of X and distances between subsets of X in the standard fashion, namely,

$$\operatorname{diam} A \triangleq \max_{x, y \in A} d(x, y)$$

$$\operatorname{dist}(A, B) \triangleq \min_{\substack{x \in A \\ y \in B}} d(x, y)$$

where $A, B \subset X$ and $A, B \neq \phi$ (empty set).

A k-partition of X is any family $P = \{X_1, \ldots, X_k\}$ of k non empty, mutually disjoint, and exhaustive subsets of X, i.e.

$$\begin{aligned} &X_i \neq \phi \qquad 1 \leqslant i \leqslant k \\ &X_i \cap X_j = \phi \qquad i \neq j \\ &\bigcup_{i=1}^{k} X_i = X \end{aligned} \tag{1}$$

The class of all such partitions is denoted by $\mathcal{P}(k)$. To each k-partition P we attach a separation index $\alpha(k,P)$ defined by the rule

$$\alpha(k, P) = \min_{1 \leqslant q \leqslant k} \min_{\substack{1 \leqslant r \leqslant k \\ r \neq q}} \operatorname{dist}(X_q, X_r) / \max_{1 \leqslant p \leqslant k} \operatorname{diam}(X_p) \tag{2}$$

and we then say that P consists of *compact separated* (CS) *clusters* relative to d if and only if,

$$\alpha(k,P) > 1.$$

If we put

$$\bar{\alpha}(k) = \max_{P} \alpha(k, P)$$

then it follows that X can be partitioned into k CS clusters if and only if

$$\bar{\alpha}(k) > 1.$$

The following result shows that X can be partitioned into k CS clusters *in at most one way*.

Theorem 1. There is at most one k-partition P of X for which $\alpha(k,P) > 1$.

Proof. Let $P^A = \{X_1{}^A, \ldots, X_k{}^A\}$ and $P^B = \{X_1{}^B, \ldots, X_k{}^B\}$ denote two arbitrary distinct k-partitions of X. Then at least one subset in P^A, say $X_\ell{}^A$, overlaps two (or more) distinct subsets of P^B, say $X_m{}^B$ and $X_n{}^B$. It follows that

$$\operatorname{diam} X_\ell{}^A \geqslant \operatorname{dist}(X_m{}^B, X_n{}^B)$$

consequently,

$$\max_{1 \leqslant p \leqslant k} \operatorname{diam} X_p{}^A \geqslant \min_{1 \leqslant q \leqslant k} \min_{\substack{1 \leqslant r \leqslant k \\ r \neq q}} \operatorname{dist}(X_q{}^B, X_r{}^B) \tag{3}$$

In exactly the same way we obtain

$$\max_{1 \leqslant p \leqslant k} \operatorname{diam}(X_p{}^B) \geqslant \min_{1 \leqslant q \leqslant k} \min_{\substack{1 \leqslant r \leqslant k \\ r \neq q}} \operatorname{dist}(X_q{}^A, X_r{}^A) \tag{4}$$

In view of (3) and (4) we therefore have

$$\frac{1}{\alpha(k, P_A)} \geqslant \alpha(k, P_B)$$

Thus,

$$\alpha(k, P_A) > 1 \;\Rightarrow\; \alpha(k, P_B) < 1$$

Since P_B is arbitrary, we conclude that there is at most one P such that $\alpha(k,P) > 1$.

(Q.E.D.)

In general, any partition P' in $\mathcal{P}(k)$ which satisfies,

$$\alpha(k, P') = \bar{\alpha}(k)$$

will be called a maximally separated k-partition of X, relative to the index α. In view of Theorem 1 there is precisely one maximally separated partition P' when $\bar{\alpha} > 1$.

A somewhat more demanding separation index β is obtained if the subsets X_r appearing in (2) are replaced by their convex hulls CoX_r (= smallest closed convex subset of V containing X_r), i.e.,

$$\beta(k, P) = \min_{1 \leqslant q \leqslant k} \min_{\substack{1 \leqslant r \leqslant k \\ r \neq q}} \operatorname{dist}(X_q, CoX_r) / \max_{1 \leqslant p \leqslant k} \operatorname{diam}(X_p) \tag{5}$$

The relationship between α and β is established in the following.

Theorem 2. For all P in $\mathcal{P}$ (k),

$$\alpha(k, P) \geqslant \beta(k, P) \geqslant \alpha(k, P) - 1 \tag{6}$$

Proof. Since Co$X_r \supset X_r$, it follows that dist$(X_q,X_r) \geqslant$ dist$(X_q,$Co$X_r)$, consequently $\alpha(k,P) \geqslant \beta(k,P)$.

To establish the right side of the inequality (6), we first note that every element y in the convex hull of $X_r \subset X$ is a convex combination of the elements of X_r, i.e.,

$$y = \sum_{\xi \epsilon X_r} \sigma_\xi \cdot \xi$$

for some (finite) set of real numbers σ_ξ satisfying

$$0 \leqslant \sigma_\xi \leqslant 1 \qquad \text{all} \qquad \xi \epsilon X_r$$

$$\sum_{\xi \epsilon X_r} \sigma_\xi = 1.$$

If x, y are arbitrary elements in X_r and CoX_r respectively, we then have

$$\begin{aligned} d(x,y) &= \|x - y\| \\ &= \left\| x - \sum_{\xi \epsilon X_r} \sigma_\xi \xi \right\| \\ &= \left\| \sum_{\xi \epsilon X_r} \sigma_\xi (x - \xi) \right\| \\ &\leqslant \left(\sum_{\xi \epsilon X_r} \sigma_\xi \right) \operatorname{diam}(X_r) = \operatorname{diam} X_r \end{aligned} \tag{7}$$

Now, if z is any element of X_q with $q \neq r$, the triangle inequality for d gives

$$d(x,y) + d(y,z) \geqslant d(x,z)$$

or

$$d(y,z) \geqslant d(x,z) - d(x,y)$$

In view of (7) we therefore have

$$\begin{aligned} d(y,z) &\geqslant d(x,z) - \operatorname{diam}(X_r) \\ &\geqslant d(x,z) - \max_{1 \leqslant p \leqslant k} \operatorname{diam}(X_p) \end{aligned} \tag{8}$$

Since x, y, z are arbitrary elements of X_r, CoX_r, and X_q respectively it follows from (8) that

$$\operatorname{dist}(X_q, CoX_r) \geqslant \operatorname{dist}(X_q, X_r) - \max_{1 \leqslant p \leqslant k} \operatorname{diam}(X_p)$$

Consequently,

$$\min_{1\leqslant q\leqslant k}\ \min_{\substack{1\leqslant r\leqslant k\\ r\neq q}} \operatorname{dist}(X_q, CoX_r) \geqslant \min_{1\leqslant q\leqslant k}\ \min_{\substack{1\leqslant r\leqslant k\\ r\neq q}} \operatorname{dist}(X_q, X_r) - \max_{1\leqslant p\leqslant k} \operatorname{diam}(X_p)$$

and therefore

$$\beta(k,P) \geqslant \alpha(k,P) - 1. \qquad \text{(Q.E.D.)}$$

The parameter β was introduced in [1], principally because of its explicit connection with fixed points of the clustering algorithm known as ISODATA. This algorithm generates partitions which are local extrema for the standard k-means least squared error criterion function, and it was shown in [1] that $\beta(k,P) > 1$ implies that P is a fixed point of the ISODATA process. The distinction between α and β is most significant when $\alpha \approx 1$, vis a vis the relationship between maximally separated partitions of X and partitions which minimize the squared error criterion or the fuzzy extensions of this criterion considered in the next section. This distinction becomes less and less important as β increases beyond 1. From Theorem 2, we can see at once that α is large compared to 1 if and only if β is large compared to 1. Furthermore, since $\beta(k,P) > 1$ implies that $\alpha(k,P) > 1$ it follows from Theorem 1 that there is at most one partition P for which $\beta(k,P) > 1$. In fact, if we put

$$\overline{\beta}(k) = \max_{p\in\wp(k)} \beta(k,P)$$

and say that P' is maximally separated relative to β if and only if

$$\beta(k,P') = \overline{\beta}(k)$$

then it follows readily from Theorems 1 and 2 that there is precisely one maximally separated partition P' relative to β when $\overline{\beta} > 1$ and that this partition P' is also maximally separated relative to α, i.e., $\alpha(k,P')=\overline{\alpha}(k)$. Since $\overline{\alpha}>2$ implies that $\overline{\beta}>1$, we can now see that the distinction between CS and CWS clusters is useful only when $2 \geqslant \alpha > 1$ and $1 \geqslant \overline{\beta}$.

3. Optimal Fuzzy Partitions and Maximally Separated Partitions

Let $u_i(\cdot):X \to \{0,1\}$ denote the characteristic function of the i^{th} subset X_i in a k-partition P, i.e.,

$$u_i(x) = \begin{cases} 0 & ;\ x \notin X_i \\ 1 & ;\ x \in X_i \end{cases}$$

and let $u(\cdot)$ denote the vector valued function $\{u_1(\cdot), \ldots, u_k(\cdot)\}$. In view of (1), $u(\cdot)$ satisfies

i) $u_i(\cdot) \not\equiv 0$

ii) $u_i(x) = 0$ or 1 for all x in X

$$\text{iii)} \quad \sum_{i=1}^{k} u_i(x) = 1 \qquad \text{for all } x \text{ in } X \tag{9}$$

Conversely, any vector valued function $u(\cdot) = \{u_1(\cdot), \ldots, u_k(\cdot)\}$ which satisfies condition (9) describes a k-partition of X, viz., the partition whose subsets X_i have $u_i(\cdot)$ as their characteristic functions. Accordingly, we will denote the class of all functions $u(\cdot)$ satisfying (9) by the same symbol $\mathcal{P}(k)$ used previously to denote the class of all k-partitions of X. With this understanding, the standard k-means least squared error partitioning problem can be stated in this way: if $\mathcal{L}$ denotes the smallest linear subspace of V continuing X and $\mathcal{L}^k$ denotes the set of all k-tuples $v = \{v_1, \ldots, v_k\}$ from $\mathcal{L}$, find a $u'(\cdot)$ in $\mathcal{P}(k)$ and a v' in $\mathcal{L}^k$ such that

$$\text{a.} \quad J(u'(\cdot), v') = \min_{u(\cdot)\in\mathcal{P}(k)} \min_{v\in\mathcal{L}^k} J(u(\cdot), v) \tag{10}$$

where

$$\text{b.} \quad J(u(\cdot), v) \triangleq \sum_{i=1}^{k} \sum_{x\in X} u_i(x)\|x - v_i\|^2 = \sum_{i=1}^{k} \sum_{x\in X} u_i(x) \langle x - v_i | x - v_i\rangle$$

For reasons discussed at length in [1] there are inherent advantages in considering various fuzzy extensions of the foregoing k-means partitioning problem. An infinite variety of smooth extensions of this kind are possible. For example, let $\mathcal{P}_f(k)$ denote the class of all fuzzy partitions of X, i.e., all functions $u(\cdot) = \{u_1(\cdot), \ldots, u_k(\cdot)\}$ which satisfy

$$\begin{aligned} &\text{i)} \quad u_i(\cdot) \not\equiv 0 \\ &\text{ii)} \quad 0 \leqslant u_i(x) \leqslant 1 \qquad \text{for all } x \text{ in } X \\ &\text{iii)} \quad \Sigma u_i(x) = 1 \qquad \text{for all } x \text{ in } X \end{aligned} \tag{11}$$

Then for any real number $\omega \geqslant 1$,

$$J_\omega(u(\cdot), v) = \sum_{i=1}^{k} \sum_{x\in X} (u_i(x))^\omega \|x - v_i\|^2 \tag{12}$$

defines a continuous and differentiable extension of J from $\mathcal{P}(k) \times \mathcal{L}^k$ to $\mathcal{P}_f(k) \times \mathcal{L}^k$. For all such extensions it is plausible that when $\bar{\alpha}$ is sufficiently large, the component membership functions in an optimal fuzzy partition, $u'(\cdot)$, satisfying

$$J_\omega(u'(\cdot), v') = \min_{u(\cdot)\in\mathcal{P}_f(k)} \min_{v\in\mathcal{L}^k} J_\omega(u(\cdot), v) \tag{13}$$

should closely approximate the characteristic functions of the subsets X_i' in the maximally separated partition P' solving,

$$\alpha(k, P') = \bar{\alpha}(k).$$

Furthermore, when $\alpha \approx 1$ or smaller it is plausible that the values of the optimal membership functions $u_i'(\cdot)$ may depart significantly from the hard limits 0 or 1 over certain regions of X. Several examples substantiating the first conjecture were given in [1] for the cases $\omega = 1$ and 2, and the second conjecture for $\omega = 2$[1]. We now give a rigorous general proof of the first conjecture after first establishing the following necessary preliminary results.

Theorem 3. $(u'(\cdot), v')$ solves (13), only if

$$\left[\sum_{x \in X} (u_i'(x))^{\omega}\right] v_i' = \sum_{x \in X} (u_i'(x))^{\omega} x. \tag{14}$$

Proof. By a straightforward generalization of the proofs for $\omega = 1$ and 2 given in [1]. Details are omitted in the interest of brevity.

Theorem 4.[2] If $(u'(\cdot), v')$ solves (13), then

$$J_\omega(u'(\cdot), v') = \frac{1}{2} \sum_{i=1}^{k} \frac{1}{\sum_{x \in X} (u_i'(x))^{\omega}} \sum_{x \in X} \sum_{y \in X} [u_i'(x) u_i'(y)]^{\omega} \|x - y\|^2 \tag{15}$$

Proof. We have

$$[u_i'(x) u_i'(y)]^{\omega} \|x - y\|^2 = [u_i'(x) u_i'(y)]^{\omega} [\langle x | x - y \rangle + \langle y | y - x \rangle]$$

therefore, with reference to (14)

$$\sum_{x \in X} \sum_{y \in X} [u_i'(x) u_i'(y)]^{\omega} \|x - y\|^2 = 2 \sum_{x \in X} [u_i'(x)]^{\omega} \sum_{x \in X} [u_i'(x)]^{\omega} \langle x | x - v_i' \rangle$$

$$= 2 \sum_{x \in X} [u_i'(x)]^{\omega} \sum_{x \in X} [u_i'(x)]^{\omega} \|x - v_i'\|^2. \tag{16}$$

Equation (15) now follows at once from (16) and (12).

Theorem 5. Let $P = \{X_1, \ldots, X_k\}$ denote an arbitrary k-partition of X, let $u'(\cdot)$ denote an optimal fuzzy partition for J_ω, and let n = number of elements in X. Then,

$$\frac{2n}{[\alpha(k, P)]^2} \geqslant \sum_{i=1}^{k} \min_{1 \leqslant q \leqslant k} \left\{ \sum_{x \in X - X_q} (u_i'(x))^{\omega} \right\}. \tag{17}$$

Proof. Let $u_i(\cdot)$ denote the characteristic function of $X_i \in P$, let $u(\cdot) = \{u_1(\cdot), \ldots, u_k(\cdot)\}$, let $v_i = \bar{x}_i$ = centroid of X_i, and let $v = \{v_1, \ldots, v_k\}$. Then $u(\cdot) \in \wp(k) \subset \wp_f(k)$, $v \in \mathcal{L}^k$, and we therefore have:

[1]When $\omega = 1$, the optimal fuzzy partition typically consists of characteristic functions (i.e., $u'(\cdot) \in \wp(k) \subset \wp_f(k)$) for *all* values of $\bar{\alpha}$; in other words, the solutions of the relaxed minimization problem (13) typically coincide with the solutions of the original minimization problem (10), regardless of whether $\bar{\alpha}$ is large or small. For this reason, the extension (12) corresponding to $\omega = 1$ is not useful from the standpoint of *detecting* the presence or absence of separated clusters in X. This point is discussed at length in [1].

[2]An analogous result holds for the standard k-means criterion in (10); cf. [3].

$$J_\omega(u'(\cdot), v') \leqslant J_\omega(u(\cdot), v)$$

$$= \sum_{i=1}^{k} \sum_{x \in X} [u_i(x)]^\omega \|x - v_i\|^2$$

$$= \sum_{i=1}^{k} \sum_{X_i} \|x - \bar{x}_i\|^2 .$$

Since $\bar{x}_i \epsilon \mathrm{Co} X_i$, it follows that

$$\|x - \bar{x}_i\|^2 \leqslant [\mathrm{diam}\,(X_i)]^2$$

(see eq. (7) in the proof of Theorem 2). Consequently, if n_i = number of elements in X_i, we have,

$$J_\omega(u'(\cdot), v') \leqslant \sum_{i=1}^{k} n_i [\mathrm{diam}\,(X_i)]^2$$

and therefore

$$J_\omega(u'(\cdot), v') \leqslant n \max_{1 \leqslant p \leqslant k} [\mathrm{diam}\,(X_p)]^2$$

$$= n \left[\max_{1 \leqslant p \leqslant k} \mathrm{diam}\,(X_p) \right]^2 . \tag{18}$$

On the other hand, by Theorem 4, we have

$$J_\omega(u'(\cdot), v') = \frac{1}{2} \sum_{i=1}^{k} \frac{1}{\sum_{x \in X} (u_i'(x))^\omega} \sum_{q=1}^{k} \sum_{r=1}^{k} \sum_{x \in X_q} \sum_{y \in X_r} [u_i'(x) u_i'(y)]^\omega \|x - y\|^2$$

$$\geqslant \frac{1}{2} \sum_{i=1}^{k} \frac{1}{\sum_{x \in X} (u_i'(x))^\omega} \sum_{q=1}^{k} \sum_{\substack{r=1 \\ r \neq q}}^{k} \sum_{x \in X_q} \sum_{y \in X_r} [u_i'(x) u_i'(y)]^\omega [\mathrm{dist}(X_q, X_r)]^2$$

$$\geqslant \frac{1}{2} \left\{ \sum_{i=1}^{k} \min_{1 \leqslant q \leqslant k} \sum_{x \in X - X_q} (u_i'(x))^\omega \right\} \cdot \min_{1 \leqslant q \leqslant k} \min_{\substack{1 \leqslant r \leqslant k \\ r \neq q}} [\mathrm{dist}(X_q, X_r)]^2 . \tag{19}$$

The inequality (17) now follows at once from (18), (19), and (2).

(Q.E.D.)

Corollary. If $P' = \{X_1', \ldots, X_k'\}$ is a maximally separated partition relative to α, then for $1 \leqslant i \leqslant k$

$$0 \leqslant \min_{1 \leqslant q \leqslant k} \sum_{X - X_q} (u_i'(x))^{\omega} \leqslant \frac{2n}{[\bar{\alpha}(k)]^2} . \tag{20}$$

As an immediate consequence of (20), we have for $1 \leqslant i \leqslant k$,

$$x \notin X'_{\gamma(i)} \Rightarrow 0 \leqslant u_i'(x) \leqslant \left(\frac{2n}{[\bar{\alpha}(k)]^2}\right)^{1/\omega} . \tag{21}$$

where $\gamma(i)$ satisfies

$$\sum_{X - X'_{\gamma(i)}} (u_i'(x))^{\omega} = \min_{1 \leqslant q \leqslant k} \sum_{X - X'_q} (u_i'(x))^{\omega} \tag{22}$$

Although the upper bound in (21) is coarse, it clearly shows that $u_i(\cdot)$ can be made arbitrarily small on $X - X'_{\gamma(i)}$, by taking $[\bar{\alpha}(k)]$ to be sufficiently large; it also suggests $u_i(\cdot)$ tends to toward 1 on $X'_{\gamma(i)}$ with increasing $\bar{\alpha}$, provided γ is a one-to-one mapping of $\{1, \ldots, k\}$ onto itself. In order to establish this last condition, we require the following simple.

Lemma 1. Let $P = \{X_1, \ldots, X_k\}$ be any family of subsets of X and let γ be any mapping of $\{1, \ldots, k\}$ into itself. Then

$$\bigcap_{i=1}^{k} (X - X_{\gamma(i)}) = X - \bigcup_{i=1}^{k} X_{\gamma(i)}. \tag{23}$$

Proof. $X - X_{\gamma(i)} \triangleq \bar{X}_{\gamma(i)}$ = complement of $X_{\gamma(i)}$ in X. Therefore

$$\bigcap_{i=1}^{k} (X - X_{\gamma(i)}) = \bigcap_{i=1}^{k} \bar{X}_{\gamma(i)}$$

$$= \overline{\left(\bigcup_{i=1}^{k} X_{\gamma(i)}\right)}$$

$$= X - \bigcup_{i=1}^{k} X_{\gamma(i)}. \qquad \text{(Q.E.D.)}$$

Theorem 6. Let ϵ denote an arbitrary positive number smaller than 1, i.e.,

$$0 < \epsilon < 1$$

and let $\bar{\alpha}$ be so large that for $1 \leqslant i \leqslant k$

$$x \notin X'_{\gamma(i)} \Rightarrow 0 \leqslant u_i'(x) \leqslant \frac{\epsilon}{k} \tag{24}$$

where $X'_{\gamma(i)} \epsilon P'$ = maximally separated k-partition relative to α [see (21)] and where γ satisfies condition (22). Then γ is one-to-one, and for $1 \leqslant i \leqslant k$,

$$x \epsilon X'_{\gamma(i)} \Rightarrow 1 - \frac{(k-1)}{k} \epsilon \qquad u_i(x) \leqslant 1. \tag{25}$$

Proof. Suppose γ is not one-to-one; then $\bigcup_{i=1}^{k} X'_{\gamma(i)}$ does not exhaust X. Therefore by Lemma 1,

$$\bigcap_{i=1}^{k} (X - X'_{\gamma(i)}) \neq \phi(\text{empty set}).$$

It follows from (24) that for some $x^* \epsilon \bigcap_{i=1}^{k} (X - X'_{\gamma(i)})$,

$$\sum_{i=1}^{k} u'_i(x^*) < \sum_{i=1}^{k} \frac{\epsilon}{k} < \epsilon < 1$$

which contradicts condition (11)iii. Thus γ is one-to-one and we therefore have for $1 \leqslant i \leqslant k$,

$$\bigcap_{\substack{\ell=1 \\ \ell \neq i}}^{k} (X - X'_{\gamma(\ell)}) = X - \bigcup_{\substack{\ell=1 \\ \ell \neq i}}^{k} X'_{\gamma(\ell)} = X'_{\gamma(i)}.$$

In view of (24) and (11)iii, it follows that for $1 \leqslant i \leqslant k$,

$$x \epsilon X'_{\gamma(i)} \Rightarrow 1 \geqslant u'_i(x) = 1 - \sum_{\substack{\ell=1 \\ \ell \neq i}}^{k} u'_\ell(x)$$

$$\geqslant 1 - \left(\frac{k-1}{k}\right)\epsilon. \qquad \text{(Q.E.D.)}$$

Theorem 6 shows that the membership functions $u'_i(\cdot)$ in an optimal fuzzy partition for (13) will approximate the characteristic functions of some permutation of the subsets X'_i in a maximally separated k-partition P', with arbitrarily high accuracy as $\overline{\alpha}(k)$ increases without bound.

References

[1] J.C. Dunn, "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters", J. Cybern., July-Sept., 1973.
[2] C.W. Curtis, *Linear Algebra*, Boston: Allyn and Bacon, Inc., 1963.
[3] R.O. Duda, and P.E. Hart, *Pattern Classification and Scene Analysis*, New York: Academic Press, 1972.

Received September, 1973