Developments in Statistics
Andrej Mrvar and Anuška Ferligoj (Editors)
Metodološki zvezki, 17, Ljubljana: FDV, 2002

A View of Some Centrality and Consensus Functions in Classification Theory and Beyond

Fred R. McMorris[1]

Abstract

The notions of centrality and distance-based consensus are important concerns in many areas such as social network theory and classification theory. The general set-up consists of a finite metric space X and a subset S of X. For $x \in X$, let $D(x, S)$ be a measure of 'remoteness' of x to S, and let C be the function where $C(S)$ is the set of all points $x \in X$ for which $D(x, S)$ is minimum. C is called the *median function* on X when $D(x, S)$ is the sum of distances from x to all the points in S, C is called the *mean function* on X when $D(x, S)$ is the sum of the squared distances, and C is called the *center function* on X when $D(x, S)$ is the maximum of the distances from x to all the points in S. This paper will review recent results obtained toward characterizing the median, mean and center functions on metric spaces such as certain classes of graphs (symmetric networks) and spaces of various types of classifications on a fixed set of entities.

1 Introduction

Let (X, d) be a finite metric space and $X^* = \bigcup_{k>1} X^k$. The elements of X^* are called *profiles* and are denoted $\pi = (x_1, \ldots, x_k)$, $\pi' = (y_1, \ldots, y_m)$, etc. In consensus theory and location theory, it is common to try to find those points in the space that are, in some sense, closest to a given profile. The literature is quite extensive, but for a start the reader can consult several of the references at the end of this note.

[1] Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616, U.S.A.
Research supported from the United States Office of Naval Research grant N00014-00-1-0004. This paper is based on a talk that was to be presented to the International Conference on Methodology and Statistics held in Ljubljana, Slovenia on 17-19 September 2001. However, the attack on the World Trade Center prevented the author from attending. Nevertheless, the author thanks the organizers, in particular Anuška Ferligoj, for the invitation.

To model the general case, we first let $D : X \times X^* \longrightarrow Re$ where Re denotes the real numbers. For appropriate functions D, $D(x, \pi)$ can be thought of as a measure of distance, or remoteness, of $x \in X$ to the profile π. In this paper, a (*distance based) consensus function* (also known as a *location function*) on X will be a function of the form $C : X^* \longrightarrow 2^X \backslash \{\emptyset\}$, where $C(\pi) = \{x : D(x, \pi) \text{ is minimum}\}$. ($2^X$ denotes the set of all subsets of X.) Let $\pi = (x_1, \ldots, x_k)$ be a profile and $x \in X$. Three natural measures of remoteness are $D_1(x, \pi) = \sum_{i=1}^{k} d(x, x_i)$, $D_2(x, \pi) = \max\{d(x_i, x) : i = 1, \ldots, k\}$, and $D_3(x, \pi) = \sum_{i=1}^{k} d(x, x_i)^2$. Current terminology (McMorris, Roberts, and Wang, 2001) is as follows: The *median function* is that consensus function, Med, where $Med(\pi) = \{x : D_1(x, \pi) \text{ is minimum}\}$, the *center function* is the function Cen where $Cen(\pi) = \{x : D_2(x, \pi) \text{ is minimum}\}$, while the *mean function*, Mea, is defined by $Mea(\pi) = \{x : D_3(x, \pi) \text{ is minimum}\}$. The most widely studied consensus function to date is the median function. However, there has been some recent progress concerning the center and mean functions, and this will discussed in the following sections. Other remoteness measures and distances would lead to other consensus functions - - - there are many possibilities to explore.

The next sections will briefly review results on the median, center, and mean functions defined on an arbitrary metric space, and then we will consider metric spaces that have additional structure that can be exploited for characterizations of Med, Cen, and Mea. The present paper extends and modifies much of what appeared in McMorris (1997).

2 The center, mean, and median functions on finite metric spaces

We first list a few properties (axioms) that a consensus function might satisfy on an arbitrary finite metric space (X, d). Each property has a straightforward interpretation for both consensus and location applications, and on the surface, each looks quite reasonable.

Consider the following properties for a consensus function C on X:

Anonymity (A): For every profile $\pi = (x_1, \ldots x_k) \in X^*$ and permutation σ of $\{1, \ldots k\}$, $C(\pi) = C(\pi^\sigma)$, where $\pi^\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(k)})$.

Betweenness (B): $C((x, y)) = \{z : d(x, y) = d(x, z) + d(z, y)\}$.

Consistency (C): If $C(\pi) \cap C(\pi') \neq \varnothing$ for profiles π and π', then $C(\pi\pi') = C(\pi) \cap C(\pi')$, where $\pi\pi'$ is the concatenation of π and π', i.e. if $\pi = (x_1, \ldots, x_k)$ and $\pi' = (y_1, \ldots, y_m)$ then $\pi\pi' = (x_1, \ldots, x_k, y_1, \ldots, y_m)$.

Faithfulness (F): $C((x)) = \{x\}$ for all $x \in X$.

If C satisfies (B) and (C), then it satisfies (F). To see this, first note that $C((x)) = C((x))$ so by consistency (C) we have $C((x,x)) = C((x))$, and $C((x,x)) = \{x\}$ by (B) since $d(x,y) = 0$ if and only if $x = y$.

It is also not hard to show that Med satisfies (A), (B) and (C) on X (and hence also (F)) (Barthélemy and Janowitz, 1991; McMorris, Mulder, and Powers, 2000). However, these properties do not in general characterize Med among consensus functions and in order to do so requires structure to be added to the metric space. More will be said about this in later sections. It is an interesting open problem to find those conditions that need to be imposed on X in order that the above three very simple properties are also sufficient and thus characterize Med.

Regarding Cen and Mea, there exist metric spaces on which neither satisfy (B) (McMorris, Roberts, and Wang, 2001). Also, it can easily be shown that Mea satisfies (C), but that Cen need not (McMorris, Roberts, and Wang, 2001). Clearly Cen and Mea both satisfy (A) and (F).

Other properties that help to separate these three consensus functions are:

Quasi-Consistency (QC): If $C(\pi) = C(\pi')$ for profiles π and π', then $C(\pi\pi') = C(\pi)$.

Population Invariant (PI): If $\{\pi\} = \{\pi'\}$, then $C(\pi) = C(\pi')$, where $\{\pi\}$ is the set of points making up the profile π.

Clearly if C satisfies (C), then it satisfies (QC), so that both Med and Mea satisfy (QC). Also, if C satisfies (PI), then C satisfies (A). Although, Cen need not satisfy (C) it does satisfy (QC) as seen in Theorem 1. Cen also obviously satisfies (PI) and thus (A), but neither Med nor Mea satisfy (PI) in general. So that this paper does not become entirely "proofless" the next result with proof from McMorris, Roberts, and Wang (2001) is included in the hope that it gives the reader some additional feeling for distance based consensus functions.

Theorem 1. The center function satisfies (QC) on any finite metric space.

Proof: For a profile $\pi = (x_1, \ldots, x_k)$ let $D(x,\pi) = \max\{d(x,x_1), \ldots, d(x,x_k)\}$. Assume that $Cen(\pi) = Cen(\pi')$, $a \in Cen(\pi\pi')$, and $b \in Cen(\pi) = Cen(\pi')$. From the definition of Cen we have $D(a,\pi\pi') \le D(b,\pi\pi'), D(b,\pi) \le D(a,\pi)$, and $D(b,\pi') \le D(a,\pi')$. Clearly

$$\begin{aligned} D(b,\pi\pi') = \max\{D(b,\pi), D(b,\pi')\} &\le \max\{D(a,\pi), D(a,\pi')\} \\ &= D(a,\pi\pi') \le D(b,\pi\pi'). \end{aligned} \tag{2.1}$$

Thus $D(a,\pi\pi') = D(b,\pi\pi')$ which implies $b \in Cen(\pi\pi')$ and therefore $Cen(\pi) \subseteq Cen(\pi\pi')$. We now claim that $a \in Cen(\pi) = Cen(\pi')$. For this, it suffices to show that $D(a,\pi) = D(b,\pi)$ or $D(a,\pi') = D(b,\pi')$. From (1), $\max\{D(b,\pi), D(b,\pi')\} = \max\{D(a,\pi), D(a,\pi')\}$. If $D(b,\pi) < D(a,\pi)$ and $D(b,\pi') < D(a,\pi')$, then $\max\{D(b,\pi), D(b,\pi')\} < \max\{D(a,\pi), D(a,\pi')\}$, a contradiction. □

3 Consensus on trees

Let $G = (X, E)$ be a finite connected graph and d be the usual geodesic metric on G, where $d(x, y)$ is the length of a shortest path between the vertices x and y. In this case a consensus function on X is referred to as a consensus function on G. Note that a profile in a graph can be thought of as a sequence of vertices, repetitions allowed. To date, there has been some success in characterizing Med on various types of graphs (Barthélemy and Janowitz, 1991; Barthélemy and McMorris, 1986; Cook and Kress, 1992; Kemeny and Snell, 1962; Leclerc, 1993; Leclerc, 1994; Margush, 1982). However, there is not much known for Cen on graphs. Cen has only been characterized on trees, where a *tree* is a finite connected graph without cycles, so results for Med will only be presented below for a tree in order to be able to compare and contrast in a straightforward fashion. In what follows $T = (X, E)$ will denote a tree. As a consequence of results holding for graphs more general than trees, the following result was established in (Leclerc, 1994).

Theorem 2. Let C be a consensus function on a tree T. Then C is the median function Med on T if and only if C satisfies properties (A), (B) and (C).

Let x and y be vertices of T and let p be the path in T joining x and y. By repeatedly deleting the end vertices of paths, starting with p, either we end with a single vertex or two adjacent vertices. We call this vertex, or two adjacent vertices, the *middle* of the path p. The following axiom is easily seen to be satisfied by Cen on T.

Middleness (M): Let C be a consensus function on a tree T, and x, y vertices in T not necessarily distinct. Then $C((x, y))$ is the middle of the unique path joining x and y in T.

In light of Theorem 2, it is tempting to conjecture that (M), (PI) and (QC) are enough to characterize Cen on trees. Unfortunately, this is not true and we need an additional axiom. If $S \subseteq T$, we let $T(S)$ denote the smallest subtree of T containing S. For the profile π and $x \in X$, let $\pi - x$ denote the profile obtained from π by removing x everywhere in π and reducing the length accordingly. Note that if $x \notin \{\pi\}$ we have $\pi - x = \pi$. $\{\pi\} - x$ will be the set of vertices appearing in π not equal to x. The following axiom, like (M), is formulated only for trees.

Redundancy (R): Let C be a consensus function on a tree T. If x is a vertex of $T(\{\pi\} - x)$, then $C(\pi - x) = C(\pi)$.

Theorem 3. (McMorris, Roberts, and Wang, 2001) Let C be a consensus function on a tree T. Then C is the center function Cen on T if and only if C satisfies properties (M), (PI), (QC) and (R).

An open problem that remains is to give a characterization of Mea on trees. Little is known, beyond what has been already pointed out as necessary conditions, for Mea on any particular class of finite metric spaces.

4 Consensus for hierarchies

Often each point in a particular space is itself an entity with structure. This in turn allows for structure to be given to the metric space in natural ways. We look first at the case where each point is a classification of a finite set of entities S. Features of every classification scheme of S usually involve the notion of a "cluster", where the clusters are constructed so that objects of the same cluster are more similar to each other than to objects of another cluster. Thus a classification on S is just a set of nonempty subsets of S. In addition, if we let H be a given classification of S, we require $S \in H$ as well as $\{x\} \in H$ for all $x \in S$. Since a *hypergraph* on S is defined as a set of nonempty subsets of S, a classification scheme is foremost a type of hypergraph. We refer to elements $A \in H$ as *clusters* of the hypergraph H. Let $\mathcal{H}$ denote the set of all hypergraphs on S of the above type and $n = |S|$ denote the number of elements of the set S. Let d be the metric on $\mathcal{H}$ defined by

$$d(H_1, H_2) = |H_1| + |H_2| - 2|H_1 \cap H_2|$$

for $H_1, H_2 \in \mathcal{H}$, where $|H_i|$ is the number of clusters in H_i. This, of course, is the well known *symmetric difference* metric. From now on the metric spaces considered will be sets of hypergraphs, equipped with the symmetric difference metric.

A classification of S will usually be structured into something tree-like called a hierarchy. In our context a *hierarchy* (also called an *n-tree*) on S is a hypergraph T, in $\mathcal{H}$, such that $A \cap B \in \{A, B, \varnothing\}$ for every cluster $A, B \in T$. Let $\mathcal{T}$ denote the set of all hierarchies on S. Because the median function on $(\mathcal{T}, d)$ is defined in terms of d, one way of "knowing" Med is to characterize this symmetric difference metric. This was done in Margush (1982) generalizing work found in Barthélemy (1979); Bogart (1973); Bogart (1975); Kemeny and Snell (1962). Also see Cook and Kness (1992) for several examples of this approach.

In order to present a characterization of Med on $(\mathcal{T}, d)$ one more property needs to be defined. It is named after the Marquis de Condorcet who, according to Young (1988), actually was suggesting the median function in 1785 (Condorcet, 1785). If $\pi = (T_1, \ldots, T_k)$ is a profile of hierarchies and $A \subseteq S$, let $\gamma(A, \pi) = \frac{|\{i: A \in T_i\}|}{k}$. A consensus function C on $(\mathcal{T}, d)$ is $\frac{1}{2}$-*condorcet* if for any $A \subseteq S$ and profile $\pi = (T_1, \ldots, T_k) \in \mathcal{T}^*$ such that $\gamma(A, \pi) = \frac{1}{2}$, the following holds:

$T \in C(\pi)$ if and only if $T \cup \{A\} \in C(\pi)$ provided $T \cup \{A\}$ is a hierarchy.

The following result appears in McMorris and Powers (1995) in an abstract version, and improved earlier results in Barthélemy and Janowitz (1991) and Barthélemy and Monjardet (1991).

Theorem 4. Let C be a consensus function on $(\mathcal{T}, d)$. Then C is the median function if and only if C satisfies properties (C) and (F) and is $\frac{1}{2}$-condorcet.

Cen and mea have not been studied on $(\mathcal{T}, \lceil)$, and this remains an interesting area for investigation.

For $\pi = (H_1, \dots, H_k) \in \mathcal{H}^*$, let $Maj\,(\pi) = \{A : \gamma(A, \pi) > \frac{1}{2}\}$. The resulting consensus function $Maj : \mathcal{H}^* \to \mathcal{H}$ is the *majority rule*. Note here that, although we refer to Maj as a consensus function, the codomain has been modified from $2^{\mathcal{H}} \backslash \{\emptyset\}$ to simply $\mathcal{H}$ - - or, one can think of the usual codomain and note that $|Maj(\pi)| = 1$ for all π. When $\mathcal{H} = \mathcal{T}$, it is easy to see that $Maj\,(\pi) \in \mathcal{T}$ for all $\pi \in \mathcal{T}^*$ (Margush and McMorris, 1981), but care must be taken for other classes of hypergraphs. For example, there exist profiles π of weak hierarchies for which $Maj\,(\pi) \notin \mathcal{W}$ (McMorris and Powers, 1991). Using Maj, finding medians in $(\mathcal{H}, d)$ is easy. Let $E(\pi) = \{A : \gamma(A, \pi) = \frac{1}{2}\}$. An argument given in Barthélemy and McMorris (1986) can be used to show that $H \in \mathcal{H}$ is a median for π if and only if $H = Maj\,(\pi) \cup K$ where $K \subseteq E(\pi)$. It follows from this that the number of medians for even a profile of two hierarchies can grow exponentially (this also follows as a consequence of axiom (B)), even though finding one median (the majority rule hierarchy) can be done in polynomial time.

5 Consensus for weak hierarchies

A hypergraph W on S is a *weak hierarchy* if and only if $A \cap B \cap C \in \{A \cap B, A \cap C, B \cap C\}$ for all clusters $A, B, C \in W$. Weak hierarchies were introduced in Bandelt and Dress (1989) to generalize hierarchies in a way that allows partial overlap of clusters. Let $\mathcal{W}$ denote the set of all weak hierarchies on S.

In contrast to the situation for hierarchies, there exist consensus functions on $(\mathcal{W}, d)$ that satisfy the three conditions of Theorem 4 yet are not the median function (McMorris and Powers, 1997). Unfortunately the major additional condition is fairly complicated. For the profile $\pi = (W_1, \dots W_k) \in \mathcal{W}^*$ and set $A \subseteq S$ define

$$w(A) = k(2\gamma(A, \pi) - 1)$$

and

$$J(\pi) = \left\{A : \gamma(A, \pi) \geq \frac{1}{2}\right\}.$$

Note that $w(A) \geq 0$ if and only if $A \in J(\pi)$. For each $W \in \mathcal{W}$ set

$$w_\pi(W) = \sum w(A)$$

where the sum is taken over all $A \in J(\pi)$ for which $A \notin W$. When $J(\pi) \subseteq W$, set $w_\pi(W) = 0$. A consensus function F on $\mathcal{W}$ is $\frac{1}{2}$-*weighted* if, for any profile $\pi \in \mathcal{W}^*$, $F(\pi) \subseteq \{W \in \mathcal{W} : w_\pi(W)$ is minimized $\}$.

In McMorris and Powers (1997) the following is proved:

Theorem 5. The median function on $(\mathcal{W}, d)$ is the unique maximum (with respect to set inclusion) consensus function on $\mathcal{W}$ which satisfies properties (C) and (F) and is $\frac{1}{2}$-condorcet and $\frac{1}{2}$-weighted.

Clearly it would be nice to find a more simple list of properties that characterize Med on $(\mathcal{W}, d)$.

We close by noting two results from McMorris and Powers (1991) that are consequences of more general theorems. The general results allow for profiles to be mixed, in the sense that members of a profile can be either hierarchies or weak hierarchies. Extreme cases of this are to consider a consensus functions $C : \mathcal{W}^* \to \mathcal{H}$ defined by $C(\pi) = \{A : \gamma(A, \pi) > \frac{1}{3}\}$ and $C : \mathcal{W}^* \to \mathcal{W}$ defined by $C(\pi) = \{A : \gamma(A, \pi) > \frac{2}{3}\}$. Both of these are clearly analogs of the majority rule in these contexts, and have been characterized while noting that both $\frac{1}{3}$ and $\frac{2}{3}$ are the smallest fractions that allow these functions to be well-defined in the sense that they produce elements of $\mathcal{H}$ and $\mathcal{W}$ respectively.

References

[1] Bandelt, H.J. and Dress, A.W.M. (1989): Weak hierarchies associated with similarity measures – An additive clustering technique. *Bull. Math. Biol.*,**43**, 133-166.

[2] Barthélemy, J.P. (1979): Caractérisations axiomatiques de la distance de la différence symetrique entre des relations binaires. *Math. Sci. hum.*, **67**, 85-113.

[3] Barthélemy, J.P. and Janowitz, M.F. (1991): A formal theory of consensus. *SIAM J. Discrete Math.*, **4**, 305-322.

[4] Barthélemy, J.P., Leclerc B., and Monjardet, B.(1986): On the use of ordered sets in problems of comparison and consensus of classifications. *J. Classification*, **3**, 187-224.

[5] Barthélemy, J.P. and McMorris, F.R. (1986): The median procedure for n-trees. *J. Classification*, **3**, 329-334.

[6] Barthélemy, J.P. and Monjardet, B. (1981): The median procedure in cluster analysis and social choice theory. *Math. Social. Sci.*, **1**, 235-268.

[7] Barthélemy, J.P. and Monjardet, B. (1988): The median procedure in sata analysis: New Results and Open Problems. In H. H. Bock (Ed.): *Classification and Related Methods of Data Analysis*, Elsevier Science, Amsterdam, 309-316.

[8] Bogart, K.P. (1973): Preference structures I: Distances between transitive preference relations. *J. Math. Sociology*, **3**, 49-67.

[9] Bogart, K.P. (1975), Preference structures II: Distances between intransitive preference relations. *SIAM J. Applied Math.*, **29**, 254-262.

[10] Condorcet, M. (1785): *Essai sur l 'Application de l 'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. Paris.

[11] Cook, W.D. and Kress, M. (1992): *Ordinal Information and Preference Structures: Decision Models and Applications*. Prentice Hall, Englewood Cliffs.

[12] Daskin, M.S. (1995): *Network and Discrete Location: Models, Algorithms, and Applications*. Wiley-Interscience, New York.

[13] Kemeny, J.G. and Snell, J.L. (1962): *Mathematical Models in the Social Sciences*. Ginn and Co., NY.

[14] Leclerc, B. (1993): Lattice valuations, medians and majorities. *Discrete Math.*, **111**, 345-356.

[15] Leclerc, B. (1994): Median for weight metrics in the covering graphs of semilattices. *Discrete Appl. Math.*, **49**, 281-297.

[16] Margush, T. (1982): Distances between trees. *Discrete Appl. Math.*, **4**, 281-290.

[17] Margush, T. and McMorris, F.R. (1981): Consensus n-trees. *Bull. Math. Biol.*, **43**, 239-244.

[18] McMorris, F.R. (1990): The median procedure for n-trees as a maximum likelihood method. *J. Classification*, **7**, 77-80.

[19] McMorris, F.R. (1997): The median function on structured metric spaces. *Student*, **2**, 195-201.

[20] McMorris, F.R., Mulder, H.M., and Powers, R.C. (2000): The median function on median hraphs and semilattices. *Discrete Appl. Math.*, **101**, 221-230.

[21] McMorris, F.R., Mulder, H.M., and Roberts, F.S. (1998): The median procedure on median graphs. *Discrete Appl. Math.*, **84**, 165-181.

[22] McMorris, F.R. and Powers, R.C. (1991): Consensus weak hierarchies. *Bull. Math. Biol.*, **53**, 679-684.

[23] McMorris, F.R. and Powers, R.C. (1995): The median procedure in a formal model of consensus, *SIAM J. Discrete Math.*, **14**, 507-516.

[24] McMorris, F.R. and Powers, R.C. (1997): The median function on weak hierarchies. To appear in B. Mirkin et al. (Eds.): *Mathematical Hierarchies and Biology*, DIMACS Series, AMS, Providence, R.I.

[25] McMorris, F.R., Roberts, F.S., and Wang, C. (2001): The center function on trees. *Networks*, **38**, 1-4.

[26] Mirchandani, P.B. and Francis R.L. (Eds.) (1990): *Discrete Location Theory.* John Wiley and Sons, NY.

[27] Monjardet, B. (1980): Théorie et application de la médiane dans les treilles distributifs finis. *Ann. Discrete Math.*, **9**, 87-91.

[28] Vohra, R. (1996): An axiomatic characterization of some locations in trees. *European J. Operational Res.*, **90**, 78-84.

[29] Young, H.P. (1988): Condorcet's theory of voting. *Am. Pol. Sci. Rev.*, **82**, 1231-1244.