Spread measures
and their relation to aggregation functions

Marek Gagolewski[a,b]

[a] *Systems Research Institute, Polish Academy of Sciences*
ul. Newelska 6, 01-447 Warsaw, Poland
[b] *Faculty of Mathematics and Information Science, Warsaw University of Technology*
ul. Koszykowa 75, 00-662 Warsaw, Poland

Abstract

The theory of aggregation most often deals with measures of central tendency. However, sometimes a very different kind of a numeric vector's synthesis into a single number is required. In this paper we introduce a class of mathematical functions which aim to measure spread or scatter of one-dimensional quantitative data. The proposed definition serves as a common, abstract framework for measures of absolute spread known from statistics, exploratory data analysis and data mining, e.g. the sample variance, standard deviation, range, interquartile range (IQR), median absolute deviation (MAD), etc. Additionally, we develop new measures of experts' opinions diversity or consensus in group decision making problems. We investigate some properties of spread measures, show how are they related to aggregation functions, and indicate their new potentially fruitful application areas.

Keywords: Group decisions and negotiations, aggregation, spread, deviation, variance

1. Introduction

Many introductory textbooks on applied statistics (or academic lectures on the subject) include a review of the so-called descriptive statistics, i.e. methods for summarizing quantitative data. Most often such methods are divided into at least two classes (cf. Aczel, 1996, Chap. 1 and e.g. Cramér, 1946):

*Email: gagolews@ibspan.waw.pl; Phone: +48 22 3810 393.

1. *measures of central tendency* (also known as measures of location or centrality of observations); e.g. sample quantiles (including median, min, and max), arithmetic mean, mode, trimmed and winsorized mean etc.,
2. *measures of variability* (or data spread), e.g. range, interquartile range, variance, standard deviation.

At the most general level, the process of combining multiple numeric values into a single, representative number is called aggregation. The theory of aggregation became a genuine, rapidly developing research field in the 1980s (see e.g. Beliakov et al., 2007; Calvo et al., 2002; Grabisch et al., 2009, 2011a,b). It may be observed, however, that the aggregation theory mainly focuses on the above-mentioned measures of central tendency, e.g. generalized means (OWA, OWMax operators, quasi-arithmetic means, etc.), averages, or "averaging functions". Such a broad class of tools is characterized by the following widely accepted definition of an aggregation function (see Grabisch et al., 2009, Def. 1.1).

Definition 1. Let $\mathbb{I} = [a, b]$. $\mathsf{A} : \mathbb{I}^n \to \mathbb{I}$ is an *aggregation function* if at least:

(a1) it is nondecreasing in each variable, i.e. for all $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$ such that $\mathbf{x} \le_n \mathbf{x}'$, i.e. $(\forall i)\ x_i \le x_i'$, it holds $\mathsf{A}(\mathbf{x}) \le \mathsf{A}(\mathbf{x}')$,

and fulfills the boundary conditions:

(a2) $\inf_{\mathbf{x}\in\mathbb{I}^n} \mathsf{A}(\mathbf{x}) = \inf \mathbb{I}$,
(a3) $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{A}(\mathbf{x}) = \sup \mathbb{I}$.

It is true that these characteristic properties reflect somehow the concept of data synthesis: finding a value representative to the whole vector. Moreover, it is well-known that such functions are strongly connected to monotone (fuzzy) measures and integrals (cf. e.g. Greco et al., 2014).

Aggregation functions have many successful applications, for example in multicriteria or group decision making, statistics, quality management, engineering, approximate reasoning, fuzzy sets and fuzzy logic (cf. the notion of a t-norm and t-conorm, which are particular aggregation functions in $[0, 1]^2$), etc.

Example 1. In a group decision making problem, assume that n decision makers express as $x_1, \ldots, x_n \in [0, 1]$ the strength of preference toward an alternative. An aggregation function may be used to combine these assessments in order to obtain a global score $\mathsf{A}(x_1, \ldots, x_n)$. For example, let $n = 4$ and $\mathbf{x} = (1, 1, 1, 0)$. If all the experts have the same standing, one may use e.g. the arithmetic mean to combine their opinions; in such a case we get $\mathsf{A}(\mathbf{x}) = 0.75$. However, assume that the fourth decision maker is conceived as less competent (at least in a given matter) than the other ones, or his/her opinion has lower significance for some other reason (see e.g. Bernasconi et al., 2014; Saaty, 1994). If e.g. a weighting vector $\mathbf{w} = (2/7, 2/7, 2/7, 1/7)$ describes the importance of the respective judges, then by calculating the weighted mean we get $\mathsf{A}'(\mathbf{x}) = 6/7 \simeq 0.86$.

It is evident that to understand the very nature of aggregation processes better, as well as to meet the practitioners' needs, we should explore new classes of methods for summarizing quantitative data. And so, the second group of measures from the above classification of descriptive statistics consists of single numbers that quantify the broadly-conceived "variability" of mathematical objects. Let us investigate it more deeply.

An important, yet not directly connected with our task, characterization of measures of entropy or uncertainty of discrete probability mass functions (represented by numeric vectors in $[0, 1]^n$ with elements summing up to 1) was proposed by Martín et al. (2001). Such a class includes e.g. the Shannon entropy and alike, cf. also (Kostal et al., 2013). Other very loosely related measures include the notion of fuzziness of a fuzzy set, cf. (Sanchez and Trillas, 2012; Weber, 1984; Zeng and Li, 2006), multidiscances (Martin and Mayor, 2011), or a probability distribution's scale parameter estimates (non-negative, translation and ratio scale invariant functions discussed by Pitman, 1939).

Among the aggregation methods of our concern, on the other hand, we may find:

1. *Measures of absolute data spread*, e.g. standard deviation, IQR, MAD. In this case, an absolute spread measure V may accompany an aggregation function A in order to state that a numeric list $\mathbf{x}$ is concisely described as $\mathsf{A}(\mathbf{x}) \pm \mathsf{V}(\mathbf{x})$.
2. *Measures of relative data spread* (e.g. Gini coefficient, coefficient of variation), which are dependent on the order of magnitude of a numeric list's elements. For instance, imagine that we have two groups of people. The first group consists of $(1, 2, 3)$-year-olds and the second one of $(101, 102, 103)$-year-olds. Intuitively, the relative spread of age in the first group is greater than that of the second group.

Most importantly, to our best knowledge none of these has been discussed from the point of view of aggregation theory. In particular, it is still unknown what characteristic properties link the measures within both groups. Note that even in statistics there are many functions which aim – at least theoretically – to be used for the mentioned purposes. Also, diverse application areas require treatment with different suitable measures. We strongly believe that the measures of absolute and relative spread are worth of deeper, separate studies. Hence, this contribution will focus on the first subclass.

Example 1 (cont'd). If all the experts are of the same esteem, we may use e.g. the sample standard deviation to assess the consistency of decision makers' opinions, refer e.g. to (Huang et al., 2013) for such an approach. However, if some form of weighting of the importance of opinions or their values is needed, then we should seek for a different kind of method for measuring the hetero/homogeneity. This, apart from measures of central tendency, could be an important, supplementary information on a numeric sequence, cf. e.g. (Ohki and Murofushi, 2012).

The paper is structured as follows. In Sec. 2 we propose a binary preorder which is further on used to determine whether a vector has no larger absolute spread than another one. Basing on this notion, in Sec. 3 we introduce the notion of a spread measure and indicate some additional properties that may be useful in particular application areas. In Sec. 4 we prove that the spread measures are naturally connected to aggregation functions. In Sec. 5 we show that the well-known descriptive statistics, like sample variance, standard deviation, interquartile range, range, median absolute deviation, and mean difference, are consistent with our definition and develop some new classes of functions which are of particular usefulness in DM tasks. Finally, Sec. 6 concludes the paper and indicates many ideas worth of deeper further studies.

2. Vectors' spread

Fix $n \in \mathbb{N}$ and let $\mathbb{I} = [a, b]$, $b > a$. From now on for each $c \in \mathbb{I}$ we denote by $(n * c)$ a sequence $(c, c, \ldots, c) \in \mathbb{I}^n$. Additionally, we assume that $[k] = \{1, \ldots, k\}$ and that whenever at least one argument is a sequence, then all arithmetic operations are properly vectorized, e.g. we have $\mathbf{x} + \mathbf{x}' = (x_1 + x'_1, \ldots, x_n + x'_n)$ and $\mathbf{x}+c = \mathbf{x}+(n*c) = (x_1+c, \ldots, x_n+c)$. In particular, Ind is a vectorized Boolean indicator function, i.e. $\mathrm{Ind}(c_1, \ldots, c_n) = (v_1, \ldots, v_1)$ with $v_i = 1$ iff logical condition c_i is true and 0 otherwise. What is more, let $x_{(i)}$, $i \in [n]$, denote the ith smallest element in $\mathbf{x} \in \mathbb{I}^n$, $\mathfrak{S}_{[n]}$ denote the set of all permutations of $[n]$, and for any $\sigma \in \mathfrak{S}_{[n]}$, $\mathbb{I}^n_\sigma = \{(x_1, \ldots, x_n) \in \mathbb{I}^n : x_{\sigma(1)} \le \cdots \le x_{\sigma(n)}\}$. Furthermore, if $\mathsf{F} : \mathbb{I}^n \to \mathbb{I}$, then let $\mathsf{F}|_\sigma$ denote the restriction of F to $\mathbb{I}^n_\sigma$, i.e. $\mathsf{F}|_\sigma : \mathbb{I}^n_\sigma \to \mathbb{I}$, $\mathsf{F}|_\sigma(\mathbf{x}) = \mathsf{F}(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{I}^n_\sigma$.

2.1. Introductory remarks

Please note that the notion of $\le_n$ plays a central role in the definition of aggregation functions. It is because an aggregation function is a morphism between the partially ordered space $(\mathbb{I}^n, \le_n)$ and linearly ordered space $(\mathbb{I}, \le)$, cf. property (a1). In other words, if $\mathbf{x} \le_n \mathbf{x}'$, then we are certain that each aggregation function ranks $\mathbf{x}$ no higher than $\mathbf{x}'$.

We shall introduce the class of absolute spread measures in a similar manner. Let us pose a question: In which case does a given vector in $\mathbb{I}^n$ *surely* have the same or not greater spread than another one in $\mathbb{I}^n$? Here is a list of the sine qua non postulates that seem reasonable for most applications.

- *Lowest possible spread.* Any constant vector, $(n * c), c \in \mathbb{I}$ should have the lowest possible spread of all the vectors considered.

- *Invariance to translations.* Spread comparison results should not change when we translate all elements in at least one sequence considered, i.e. $\mathbf{x}$ and $\mathbf{x} + t$ are of the same spread for any $\mathbf{x}, t$. Note that such a condition would be inappropriate in case of measures of relative spread.

- *Non-symmetry.* In statistics and data analysis, perhaps we will not take into account the relative ordering of the elements in a sequence: for any $\sigma', \sigma'' \in \mathfrak{S}_{[n]}$ the vectors $(x_{\sigma'(1)}, \ldots, x_{\sigma'(n)})$ and $(x_{\sigma''(1)}, \ldots, x_{\sigma''(n)})$ have the same spread, as we treat all the observations as just "points in the real line"; however, here we should be interested in a more general setting in which the relative ordering may be important: for example, each element in a vector may have a corresponding weight which is determined by its position (the ith element may be more "important" than the jth, cf. the above example).

Moreover, how to modify a given vector $\mathbf{x}$ so that its spread surely does not decrease? A sensible answer may be given in terms of the notion of some kind of distance between all the pairs of elements. Namely, if the distance between each x_i and x_j does not decrease, then the spread also does not decrease. The most natural choice of the distance measure in $\mathbb{I}$ is of course an ℓ^p-norm generated one, $d(x_i, x_j) = |x_j - x_i|$. However, according to the non-symmetry postulate, we should rather insist on checking whether the *signed* distance between each pair of observations of the first vector is not greater than the distance between the corresponding pairs from the second one, cf. (Rothschild and Stiglitz, 1970) for a well-known approach concerning increasing a spread of a probability distribution.

Please cite this paper as: Gagolewski M., Spread measures and their relation to aggregation functions,
European Journal of Operational Research **241**(2), 2015, pp. 469–477, doi:10.1016/j.ejor.2014.08.034.

2.2. Definition

The above intuitions are reflected by the following binary relation $\preccurlyeq_n$ on $\mathbb{I}^n$. Given $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$, we write $\mathbf{x} \preccurlyeq_n \mathbf{x}'$ and say that $\mathbf{x}$ *has not greater absolute spread than* $\mathbf{x}'$, if and only if for all $i, j \in [n]$ it holds:

$$(x_i - x_j)(x_i' - x_j') \geq 0 \text{ and } |x_i - x_j| \leq |x_i' - x_j'|. \tag{1}$$

Figure 1 illustrates two vectors: $\mathbf{x}$ and its modified version $\mathbf{x}'$ with increased distances between consecutive elements.

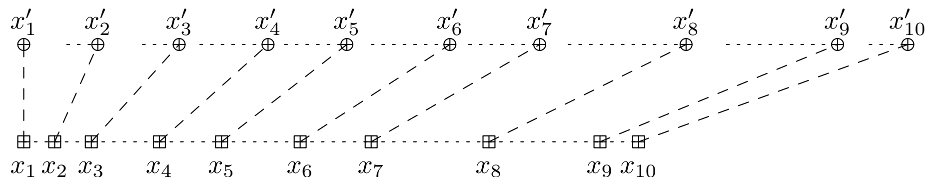

Figure 1: Two exemplary vectors with different spreads: $\mathbf{x} \preccurlyeq_n \mathbf{x}'$

Example 2. We have $(0.0, 0.1, 0.3) \preccurlyeq_n (0.0, 0.2, 0.6)$. Moreover, $(0.0, 0.1, 0.3) \preccurlyeq_n (0.2, 0.3, 0.5) \preccurlyeq_n (0.0, 0.1, 0.3)$.

2.3. Properties and examples

Please note that $\preccurlyeq_n$ is a preorder on $\mathbb{I}^n$, i.e. a relation that is reflexive and transitive. Firstly, let us study how $\preccurlyeq_n$ behaves under scaling and translation of elements in a given vector.

Remark 2. *It is easily seen that for all $s \geq 1$ and $\mathbf{x} \in \mathbb{I}^n$ such that $s\mathbf{x} = (sx_1, \ldots, sx_n) \in \mathbb{I}^n$ we have $\mathbf{x} \preccurlyeq_n s\mathbf{x}$ and of course $\frac{1}{s}\mathbf{x} \preccurlyeq_n \mathbf{x}$. Additionally, for all $t \in \mathbb{R}$ for which $t + \mathbf{x} = (t + x_1, \ldots, t + x_n) \in \mathbb{I}^n$ it holds $\mathbf{x} \preccurlyeq_n t + \mathbf{x}$ and, at the same time, $t + \mathbf{x} \preccurlyeq_n \mathbf{x}$. Thus, $\preccurlyeq_n$ is not antisymmetric.*

What is more, it is easily seen that for all $c \in \mathbb{I}$, $(n * c)$ is a minimal element of $(\mathbb{I}^n, \preccurlyeq_n)$, i.e. for any $\mathbf{x}$ we have $(n * c) \preccurlyeq_n \mathbf{x}$. This relation is also convex: for all $\mathbf{x}, \mathbf{x}', \alpha \in [0, 1]$ it holds $\mathbf{x} \preccurlyeq_n \alpha\mathbf{x} + (1 - \alpha)\mathbf{x}' \preccurlyeq_n \mathbf{x}'$ whenever $\mathbf{x} \preccurlyeq_n \mathbf{x}'$.

Note that if $\mathbb{I}$ is not a singleton, then also $\preccurlyeq_n$ is not a total (complete) binary relation: there exist $\mathbf{x}, \mathbf{x}'$ such that $\mathbf{x} \not\preccurlyeq_n \mathbf{x}'$ and $\mathbf{x}' \not\preccurlyeq_n \mathbf{x}$. In other words, not all pairs of vectors are comparable with $\preccurlyeq_n$.

Example 3. It holds $(0.0, 0.2, 0.4, 0.6, 0.8, 1.0) \not\preccurlyeq_n \not\succcurlyeq_n (0.0, 0.0, 0.0, 1.0, 1.0, 1.0)$, where of course $\succcurlyeq_n = \preccurlyeq_n^{-1}$. This fact conforms with the intuition: we cannot objectively (from all points of views) state which of the two vectors has greater spread:

the second vector represents highly "polarized" observations, but in the two subgroups $(0.0, 0.0, 0.0)$ and $(1.0, 1.0, 1.0)$ the spread is the lowest possible (we may wish to give more weight to e.g. small observations). On the other hand, the observations in the first vector are all different. We see that, in general, our relation does not indicate whether we should look at a "global" or a "local" level, which is an advantage. A similar behavior is exhibited by $\leq_n$, e.g. we have $(0.5, 0.5) \not\leq_n \not\geq_n (0, 1)$.

Additionally, whether $\preccurlyeq_n$ holds for given $\mathbf{x}, \mathbf{x}'$ depends on how the elements in both vectors are jointly ordered. The left side of (1) implies that if $\mathbf{x} \preccurlyeq_n \mathbf{x}'$, then $\mathbf{x}, \mathbf{x}'$ are *comonotonic* (cf. Grabisch et al., 2009, Def. 2.123). Thus, trivially, if $\mathbf{x} \preccurlyeq_n \mathbf{x}'$, then there exists $\sigma \in \mathfrak{S}_{[n]}$ such that $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n_\sigma$. In fact, it might be shown that σ is an ordering permutation of $\mathbf{x}'$.

Example 1 (cont'd). Our definition implies that $(1.0, 1.0, 1.0, 0.5) \not\preccurlyeq_n \not\succcurlyeq_n (0.6, 1.0, 1.0, 1.0)$. When we treat the elements as just points in the real line, the second vector has not greater spread than the first one. However, in general we cannot compare them. Such a behavior of the $\preccurlyeq_n$ relation is especially useful in group decision making. Recall that in Example 1 we had different confidence toward the competence of decision makers, and we should be really careful when judging whether they agree on a given topic or not. In our example, x_4 is known to express a less faithful judgment. Thus, in $(1.0, 1.0, 1.0, 0.5)$, the first three "faithful experts" totally agree and in $(0.8, 1.0, 1.0, 1.0)$ they do not reach full consensus.

3. Spread measures

3.1. Definition

Let us proceed with the definition of objects in which we have a special interest in this paper.

Definition 3. A *spread measure* is a mapping $\mathsf{V} : \mathbb{I}^n \to [0, \infty]$ such that:

(v1) for each $\mathbf{x} \preccurlyeq_n \mathbf{x}'$ it holds $\mathsf{V}(\mathbf{x}) \leq \mathsf{V}(\mathbf{x}')$,
(v2) for any $c \in \mathbb{I}$ it holds $\mathsf{V}(n * c) = 0$.

Note that the first characteristic property implies that each spread measure is *translation invariant*, i.e. we have:

(v0) For any $\mathbf{x} \in \mathbb{I}^n$ and $t \in \mathbb{R}$ such that $\mathbf{x} + t \in \mathbb{I}^n$ it holds $\mathsf{V}(t + \mathbf{x}) = \mathsf{V}(\mathbf{x})$.

This property is fulfilled because $(\forall t, \mathbf{x})\ \mathbf{x} \preccurlyeq_n t + \mathbf{x}$ and $t + \mathbf{x} \preccurlyeq_n \mathbf{x}$ (see Remark 2).

On the other hand, the second characteristic property gives a kind of "normalization" of a spread measure's output value: a constant vector has no spread at all. Of course, for $n = 1$, a spread measure always outputs the value 0, as a single number has "no spread". Thus, we will always be interested in the case $n \geq 2$. Also please note that for each function fulfilling (v1), by the fact that $(\forall c \in \mathbb{I})$ $(n * c) \preccurlyeq_n \mathbf{x}$, the second characteristic property is equivalent to requiring that $\inf_{\mathbf{x}\in\mathbb{I}^n} \mathsf{V}(\mathbf{x}) = 0$. We see that each V is a morphism between the preordered space $(\mathbb{I}^n, \preccurlyeq_n)$ and $([0, \infty], \leq)$, fulfilling the boundary condition (v2) concerning minima of $(\mathbb{I}^n, \preccurlyeq_n)$.

3.2. Additional properties of spread measures

Note that there is no much sense in assuming that V should be a function onto $\mathbb{I}$, because of translation invariance. 0 is a good choice for the "lowest possible spread" as we will later see that some well-known measures of spread used in statistics do indeed output such a value for $(n * c)$. On the other hand, in practice, when constructing a spread measure, one may wish to think of a kind of normalization of V by providing some condition similar to the (a3) property, for example to impose that a spread measure in $[0, 1]^n$ gives values in $[0, 1]$, or – more generally – in $[0, b - a]$. These requirements may be expressed by the conditions:

(v3) $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{V}(\mathbf{x}) = b - a$, or
(v3') $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{V}(\mathbf{x}) \leq b - a$,

In this paper (v3) and (v3') are not too important. This is because we have the following property (a similar result holds for aggregation functions).

Remark 4. *Let* $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{V}(\mathbf{x}) = u$. *Then for each nondecreasing function* $\varphi : [0, u] \to [0, (b - a)]$ *such that* $\varphi(0) = 0$, $\varphi(u) = b - a$, *if* V *is a spread measure, then* $\varphi \circ \mathsf{V}$ *is a spread measure fulfilling (v3).*

Of course, in some cases one may impose additional conditions on V, like *homogeneity of degree 1* (w.r.t. multiplication, also called *ratio scale invariance*, see Grabisch et al., 2009, Def. 2.86), cf. Remark 2:

(v4) $(\forall \mathbf{x} \in \mathbb{I}^n)$ $(\forall s > 0)$ if $s\mathbf{x} \in \mathbb{I}^n$, then $\mathsf{V}(s\mathbf{x}) = s\mathsf{V}(\mathbf{x})$,

or *continuity* (cf. Grabisch et al., 2009, Proposition 2.8):

(v5) V is continuous in each variable.

Moreover, in statistics (but perhaps not in decision making) we may assume that for each $\sigma \in \mathfrak{S}_{[n]}$ and for all vectors $\mathbf{x}$ it holds $\mathsf{V}(\mathbf{x}) = \mathsf{V}(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$, which is the same as assuming that we e.g. aggregate only nondecreasingly sorted vectors. Thus, we may need the *symmetry* property (cf. Grabisch et al., 2009, Theorem 2.34):

(v6) $(\forall \mathbf{x} \in \mathbb{I}^n)\ \mathsf{V}(\mathbf{x}) = \mathsf{V}(x_{(1)}, \ldots, x_{(n)})$,

and so on.

Example 1 (cont'd). Let $\mathbf{x} = (1, 1, 1, 0.5)$ and $\mathbf{x}' = (0.5, 1, 1, 1)$ denote the opinions of four experts on two different matters and $\mathbf{w} = (2/7, 2/7, 2/7, 1/7)$ be a weighting vector, i.e. the fourth expert is regarded as less competent than the others. The weighted means are equal to 13/14 and 12/14, respectively. The mean squared deviance from the weighted mean $\mathsf{MSD}_{\mathbf{w}}(\mathbf{y}) = \sum_{i=1}^{n}(y_i - \sum_{j=1}^{n} w_i y_i)^2/4$ gives $\mathsf{MSD}_{\mathbf{w}}(\mathbf{x}) = 39/784$ and $\mathsf{MSD}_{\mathbf{w}}(\mathbf{x}') = 37/784$. However, by using a spread measure $\mathsf{WD}_2\mathsf{WAM}_{\mathbf{w},\mathbf{w}}(\mathbf{y}) = \sum_{i=1}^{n} w_i(y_i - \sum_{j=1}^{n} w_i y_i)^2$ (see Proposition 17), we obtain $\mathsf{WD}_2\mathsf{WAM}_{\mathbf{w},\mathbf{w}}(\mathbf{x}) = 3/98$ and $\mathsf{WD}_2\mathsf{WAM}_{\mathbf{w},\mathbf{w}}(\mathbf{x}') = 5/98$, which better agrees with our intuition.

In Section 5 we will show that such functions as variance, standard deviation, range, and interquartile range (among others) are indeed spread measures. Before that, however, we should study the relationship between spread measures and aggregation functions. Additionally, we are going to explore some basic properties of spread measures.

4. Spread measures and their relation to aggregation functions

For any given $\mathbf{x} \in \mathbb{I}^n$ let $\mathsf{diff}(\mathbf{x}) = (x_{(2)} - x_{(1)}, \ldots, x_{(n)} - x_{(n-1)})$ denote the *iterated difference* between consecutive ordered components of a given vector. Please note that such a function is available in some statistical packages: for instance, it may be calculated by calling `diff(sort(x))` in R (R Development Core Team, 2014). We see that if $\boldsymbol{\delta} = \mathsf{diff}(\mathbf{x})$, then $0 \le \delta_i \le b - a$ and $\sum_{i=1}^{n-1} \delta_i \le b - a$. Intuitively, if $\mathbf{x}$ is already ordered, then this operation may be viewed as a kind of "vector differentiation". On the other hand, for $\tilde{\mathbf{x}} = \mathsf{cumsum}(x_{(1)}, \boldsymbol{\delta}) = (x_{(1)}, x_{(1)} + \delta_1, x_{(1)} + \delta_1 + \delta_2, \ldots, x_{(1)} + \delta_1 + \cdots + \delta_n)$ denoting the *cumulative sum* of $\hat{\boldsymbol{\delta}} = (x_{(1)}, \boldsymbol{\delta})$ we have $x_{(i)} = \tilde{x}_i$, $x_i = \tilde{x}_{\sigma^{-1}(i)}$, where σ is such that $\mathbf{x} \in \mathbb{I}^n_\sigma$. Thus, $\mathbf{x}$ may be reconstructed from $x_{(1)}, \boldsymbol{\delta}$, and σ.

We are now in a position to provide an equivalent definition of the relation defined by (1).

Lemma 5. *For any* $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$ *it holds* $\mathbf{x} \preccurlyeq_n \mathbf{x}'$ *if and only if* $\mathbf{x}, \mathbf{x}'$ *are comonotonic and* $\mathsf{diff}(\mathbf{x}) \leq_{n-1} \mathsf{diff}(\mathbf{x}')$.

Proof. ($\Longrightarrow$) Trivial.
($\Longleftarrow$) Take any comonotonic $\mathbf{x}, \mathbf{x}'$ and let $\sigma \in \mathfrak{S}_{[n]}$ be such that $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n_\sigma$. Also let $\boldsymbol{\delta} := \mathsf{diff}(\mathbf{x}) \leq_{n-1} \mathsf{diff}(\mathbf{x}') =: \boldsymbol{\delta}'$. Take any i, j such that $x_i \geq x_j$. Since $\boldsymbol{\delta} \leq_{n-1} \boldsymbol{\delta}'$, it holds $x_i - x_j = x_{(1)} + \sum_{k=1}^{\sigma^{-1}(i)-1} \delta_k - x_{(1)} - \sum_{k=1}^{\sigma^{-1}(j)-1} \delta_k = \sum_{k=\sigma^{-1}(j)}^{\sigma^{-1}(i)-1} \delta_k \leq \sum_{k=\sigma^{-1}(j)}^{\sigma^{-1}(i)-1} \delta'_k = x'_i - x'_j$, with convention $\sum_{k=1}^{0} \cdots = 0$. Thus, $\mathbf{x} \preccurlyeq_n \mathbf{x}'$, and the proof is complete. □

Corollary 6. *Let* $\mathcal{X} \subset \mathbb{I}^n$ *be a set such that* $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ *iff* $x_{(1)} = x'_{(1)}$. *Then for any* $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ *if* $\mathbf{x} \preccurlyeq_n \mathbf{x}'$, *then there exist* $k \geq 2$, *vectors* $\mathbf{y}^{\{1\}}, \ldots, \mathbf{y}^{\{k\}}$ *with* $\mathbf{y}^{\{1\}} = \mathbf{x}$ *and* $\mathbf{y}^{\{k\}} = \mathbf{x}'$, *constants* $d_1, \ldots, d_k > 0$ *and* $c_i, \ldots, c_k \in \mathbb{I}$ *such that* $\mathbf{y}^{\{i+1\}} = \mathbf{y}^{\{i\}} + d_i Ind(\mathbf{y}^{\{i\}} \geq c_i)$ *for all* $i \in [k-1]$. *Therefore, if* $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, *and* $\mathbf{x} \preccurlyeq_n \mathbf{x}'$, *then* $\mathbf{x} \leq_n \mathbf{x}'$ *(of course, the converse implication is not true in general).*

Remark 7. *Lemma 5 implies that checking whether* $\preccurlyeq_n$ *holds between a pair of vectors may be performed with an algorithm taking at most* $O(n \log n)$ *time: it is a matter of finding an ordering permutation of* $\mathbf{x}'$ *and some simple verifications taking linear time.*

We may now provide an alternative characterization of our class.

Theorem 8. $\mathsf{V} : \mathbb{I}^n \to [0, \infty]$ *is a spread measure if and only if*

(v1') *for each comonotonic* $\mathbf{x}, \mathbf{x}'$ *such that* $\mathsf{diff}(\mathbf{x}) \leq_{n-1} \mathsf{diff}(\mathbf{x}')$ *we have* $\mathsf{V}(\mathbf{x}) \leq \mathsf{V}(\mathbf{x}')$,

(v2') $\inf_{\mathbf{x} \in \mathbb{I}^n} \mathsf{V}(\mathbf{x}) = 0$.

This result follows directly from Lemma 5 and the fact that $\mathsf{diff}(\mathbf{x}) = (0, \ldots, 0)$ iff there exists c such that $\mathbf{x} = n * c$. It becomes particularly appealing when we recall the definition of an aggregation function and compare the conditions (v1') and (v2') to (a1) and (a2). In fact, we have what follows.

Corollary 9. *Any spread measure* $\mathsf{V} : \mathbb{I}^n \to [0, \infty]$ *may be generated by a family of functions* $\{\tilde{\mathsf{A}}_\sigma : \sigma \in \mathfrak{S}_{[n]}\}$, $(\forall \sigma \in \mathfrak{S}_{[n]})\ \tilde{\mathsf{A}}_\sigma : [0, b-a]^{n-1} \to [0, \infty]$ *fulfills (a1) and (a2), and for all* $\mathbf{x}$ *and each* σ, σ' *with* $\mathbf{x} \in \mathbb{I}^n_\sigma$ *and* $\mathbf{x} \in \mathbb{I}^n_{\sigma'}$, *it holds* $\tilde{\mathsf{A}}_\sigma(\mathbf{x}) = \tilde{\mathsf{A}}_{\sigma'}(\mathbf{x})$. *In such a setting,* $\mathsf{V}(\mathbf{x}) = \tilde{\mathsf{A}}_\sigma(\mathsf{diff}(\mathbf{x}))$ *for* $\mathbf{x} \in \mathbb{I}^n_\sigma$.

Of course, if V is symmetric, then we may restrict ourselves to the space of vectors that are already sorted. In such a case, only one generalized aggregation function is needed to generate V.

Corollary 10. *For any* $\mathsf{V} : \mathbb{I}^n \to [0, \infty]$, $\mathsf{V}|_\sigma$ *fulfills (v1) and (v2) if and only if there exists* $\tilde{\mathsf{A}} : [0, b-a]^{n-1} \to [0, \infty]$ *such that* $\mathsf{V}|_\sigma(\mathbf{x}) = \tilde{\mathsf{A}}(\mathsf{diff}(\mathbf{x}))$ *fulfills (a1) and (a2).*

Thus, we have established a beautiful connection between spread measures and aggregation functions.

Additionally, note that if $\mathsf{V}_1, \ldots, \mathsf{V}_k$ are spread measures and $\tilde{\mathsf{A}} : [0, \infty]^k \to [0, \infty]$ fulfills $(a1)$ and $(a2)$, then V' such that $\mathsf{V}'(\mathbf{x}) = \tilde{\mathsf{A}}(\mathsf{V}_1(\mathbf{x}), \ldots, \mathsf{V}_k(\mathbf{x}))$ is also a spread measure.

5. Classical spread measures and their generalizations

In this section we will provide an overview of descriptive statistics classically used in exploratory data analysis (all of them are symmetric), show that they are indeed spread measures according to our definition, and generalize them. This will lead to very interesting classes of functions that may be used e.g. in group decision making. Please note that an important step in most of the proofs below consists of rewriting $\mathsf{V}(\mathbf{x})$ in terms of $\tilde{\mathsf{A}}(\boldsymbol{\delta})$ with $\delta = \mathsf{diff}(\mathbf{x})$. Such a reparametrization is also important from the perspective of aggregation functions theory, as it may indicate some new, not yet explored classes of central tendency measures.

5.1. Distances between all pairs of observations

The first class of spread measures consists of sums of distances between all pairs of observations in a sample.

Proposition 11. *A function* V *defined for any* $p \geq 1$ *and* $\mathbf{x} \in \mathbb{I}^n$ *as* $\mathsf{V}(\mathbf{x}) = \sum_{i=1}^n \sum_{k=1}^n |x_i - x_k|^p$ *is a continuous and symmetric spread measure.*

Proof. It is easily seen that V is symmetric and fulfills (v2). Take any $\mathbf{x}$ and let $\boldsymbol{\delta} = \mathsf{diff}(\mathbf{x})$. Under the convention that $\sum_{i=1}^0 \cdots = 0$ we have:

$$\mathsf{V}(\mathbf{x}) = 2 \sum_{i=2}^n \sum_{k=1}^{i-1} (x_{(i)} - x_{(k)})^p = 2 \sum_{i=2}^n \sum_{k=1}^{i-1} \left(\sum_{j=1}^{i-1} \delta_j - \sum_{j=1}^{k-1} \delta_j \right)^p = 2 \sum_{i=2}^n \sum_{k=1}^{i-1} \left(\sum_{j=k}^{i-1} \delta_j \right)^p,$$

which is a nondecreasing and continuous function w.r.t. each element in $\boldsymbol{\delta}$. Thus, V is a spread measure. □

Recall that for any $\mathbf{x} \in \mathbb{I}^n$ and $s \geq 1$ such that $s\mathbf{x} \in \mathbb{I}^n$ we have $\mathbf{x} \preccurlyeq_n s\mathbf{x}$, see Remark 2. For any normalizing constant $\eta > 0$, a spread measure given by $\mathsf{V}(\mathbf{x}) = \left(\eta \sum_{i=1}^n \sum_{k=1}^n |x_i - x_k|^p\right)^{1/p}$ fulfills (v4), i.e. is homogeneous of degree 1.

Remark 12. *It might be shown that the* sample variance, $\mathsf{Var}(\mathbf{x}) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{\mathbf{x}})^2$, $\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n} x_i$, *may be expressed as* $\mathsf{Var}(\mathbf{x}) = \frac{1}{2n(n-1)}\sum_{i=1}^{n}\sum_{k=1}^{n}(x_i - x_k)^2$. *As a consequence, it is a symmetric and continuous spread measure. In this case, assuming that* $\boldsymbol{\delta} = \mathsf{diff}(\mathbf{x})$, *it holds* $\mathsf{Var}(\mathbf{x}) = \sum_{i=1}^{n-1}\sum_{k=1}^{n-1}\frac{\min\{i(n-k),k(n-i)\}}{n(n-1)}\delta_i\delta_k$. *Moreover, by Remark 4, the* standard deviation $\mathsf{SD}(\mathbf{x}) = \sqrt{\mathsf{Var}(\mathbf{x})}$ *is also a spread measure.*

We have $\sum_{i=1}^{n}\sum_{k=1}^{n}(x_i - x_k)^2 = \sum_{i=1}^{n} 2(n-1)x_{(i)}^2 - \sum_{1\le i<j\le n} 4x_{(i)}x_{(j)}$. *It might be shown that this formula is maximized e.g. at* $(\lfloor\frac{n}{2}\rfloor * b, \lceil\frac{n}{2}\rceil * a)$ *which is then equal to* $2\lfloor\frac{n}{2}\rfloor(n-\lfloor\frac{n}{2}\rfloor)(b-a)^2$. *Thus,* $\mathsf{V}(\mathbf{x}) = \sqrt{\eta\mathsf{Var}(\mathbf{x})}$ *fulfills (v1)–(v6) if* $\eta = 4\frac{n-1}{n}$ *for even* n *and* $\eta = 4\frac{n}{n+1}$ *otherwise.*

Remark 13. *The above proposition also implies that the* mean difference, *defined as* $\mathsf{MD}(\mathbf{x}) = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{k=1}^{n}|x_i - x_k|$, *is a symmetric spread measure. Here if* $\boldsymbol{\delta} = \mathsf{diff}(\mathbf{x})$, *then* $\mathsf{MD}(\mathbf{x}) = 2\sum_{i=1}^{n-1}\frac{i(n-i)}{n(n-1)}\delta_i$.

What is more, note that $\mathsf{MD}(\mathbf{x}) = \frac{2}{n(n-1)}\sum_{i=1}^{n}(2i-n-1)x_{(i)}$. *It is easily seen that it takes its maximal value e.g. at* $(\lfloor\frac{n}{2}\rfloor * b, \lceil\frac{n}{2}\rceil * a)$ *which then is equal to* $\frac{2}{n(n-1)}\lfloor\frac{n}{2}\rfloor\lceil\frac{n}{2}\rceil(b-a)$, *i.e.* $\frac{n}{2(n-1)}(b-a) \le (b-a)$ *for even* n *and* $\frac{(n+1)}{2n}(b-a) \le (b-a)$ *otherwise. From this it follows that* MD *fulfills (v3'). If we would like to require (v3) we may use these normalizing constants.*

5.2. Aggregated distance to quantiles

Interestingly, it turns out that *any* aggregation function on a specific domain (e.g. the weighted mean, OWA, OWMax operators, etc.) may be used to generate a spread measure in a quite simple way – basing on the fact that each object of our interest is translation invariant. It is Corollary 6 which implies that for any aggregation function $\mathsf{A} : [0, b-a]^n \to [0, b-a]$ a mapping $\mathsf{V}(\mathbf{x}) = \mathsf{A}(x_1 - x_{(1)}, \ldots, x_n - x_{(1)})$ is a spread measure. For example, the *sample range*, given by $\mathsf{Range}(\mathbf{x}) = x_{(n)} - x_{(1)}$, is included in this class.

On the other hand, the class of spread measures is in some sense much richer: such a measure, after a proper translation of its inputs, is not necessarily nondecreasing in each variable. For example, let $\mathbf{y} = (0, 0, 1)$ and $\mathbf{y}' = (0, 0.5, 1)$. We have $\mathbf{y} \le_n \mathbf{y}'$ and $\mathbf{y} \not\preceq_n \not\succeq_n \mathbf{y}'$. However, $\mathsf{Var}(\mathbf{y}) = \frac{1}{3} > \mathsf{V}(\mathbf{y}') = \frac{1}{4}$, thus Var does not fulfill (a1).

Nevertheless, the following generalization of $\mathsf{A}(x_1 - x_{(1)}, \ldots, x_n - x_{(1)})$ might be interesting as it leads to a proof of the fact that another set of data analysis tools indeed consists of spread measures. Let us recall the notion of a *sample quantile* of order $\alpha \in [0, 1]$. Although there are many various definitions in the literature and implementations in statistical software packages, see (Hyndman and

Fan, 1996) for a review, it is generally accepted that such an aggregation function is an OWA operator (Yager, 1988) given by:

$$\mathsf{Q}_\alpha(\mathbf{x}) = \gamma\, x_{(k)} + (1-\gamma)\, x_{(k+1)},$$

for some $\gamma = \gamma(\alpha, k) \in (0, 1]$ and $k \in \{\lfloor n\alpha \rfloor, \lfloor n\alpha + 1 \rfloor\}$. Also, three special cases should be included: with Q_0 being the sample minimum, Q_1 standing for sample maximum, and $\mathsf{Q}_{0.5}$ being equivalent to the sample median, $\mathsf{Med}(\mathbf{x}) = x_{((n+1)/2)}$ for odd n, and $\mathsf{Med}(\mathbf{x}) = 0.5x_{(n/2)} + 0.5x_{(n/2+1)}$ otherwise. Moreover, for each fixed $\mathbf{x}$ this should be a nondecreasing function of α.

Let us consider the following class of functions, constructed by aggregating deviations from sample quantiles.

Proposition 14. *For each $\alpha \in [0, 1]$ and any function $\tilde{\mathsf{A}} : [0, (b-a)]^n \to [0, \infty]$ fulfilling (a1) and (a2), the function $\mathsf{V} : \mathbb{I}^n \to [0, \infty]$ given by:*

$$\mathsf{V}(\mathbf{x}) = \tilde{\mathsf{A}}\left(|x_1 - \mathsf{Q}_\alpha(\mathbf{x})|, \dots, |x_n - \mathsf{Q}_\alpha(\mathbf{x})|\right) \tag{2}$$

is a spread measure.

Proof. Let $\gamma \in (0, 1]$ and $1 \le k < n$ be such that for all $\mathbf{x}$ we have $\mathsf{Q}_\alpha(\mathbf{x}) = \gamma\, x_{(k)} + (1-\gamma)\, x_{(k+1)}$. The function V fulfills (v2), because for all $c \in \mathbb{I}$ we have $\mathsf{Q}_\alpha(n * c) = c$ and by (a2) it holds $\mathsf{V}(n * c) = \tilde{\mathsf{A}}(0, \dots, 0) = 0$.

Consider any σ and $\mathbf{x} \in \mathbb{I}^n_\sigma$. Let $\hat{\boldsymbol{\delta}} = (x_{(1)}, \mathsf{diff}(\mathbf{x}))$ and $\mathbf{y} = \big(|x_1 - \mathsf{Q}_\alpha(\mathbf{x})|, \dots, |x_n - \mathsf{Q}_\alpha(\mathbf{x})|\big)$. Recall that for all $j \in [n]$ we have $x_j = \sum_{l=1}^{\sigma^{-1}(j)} \hat{\delta}_l$ and $x_{(j)} = \sum_{l=1}^{j} \hat{\delta}_l$. Thus:

$$y_j = \left|x_j - \mathsf{Q}_\alpha(\mathbf{x})\right| = \left|\sum_{l=1}^{\sigma^{-1}(j)} \hat{\delta}_l - \gamma \sum_{l=1}^{k} \hat{\delta}_l - (1-\gamma) \sum_{l=1}^{k+1} \hat{\delta}_l\right| = \left|\sum_{l=1}^{\sigma^{-1}(j)} \hat{\delta}_l - \sum_{l=1}^{k+1} \hat{\delta}_l + \gamma \hat{\delta}_{k+1}\right|.$$

As $\mathsf{V}(\mathbf{x}) = \tilde{\mathsf{A}}(\mathbf{y})$, by nondecreasingness of $\tilde{\mathsf{A}}$ it suffices to show that as elements in $\hat{\boldsymbol{\delta}}$ increase, any y_i does not decrease. Indeed, we have:

$$y_j = \begin{cases} \sum_{l=\sigma^{-1}(j)+1}^{k} \hat{\delta}_l + (1-\gamma)\, \hat{\delta}_{k+1} & \text{for} \quad \sigma^{-1}(j) \le k, \\ \sum_{l=k+2}^{\sigma^{-1}(j)} \hat{\delta}_l + \gamma\, \hat{\delta}_{k+1} & \text{for} \quad \sigma^{-1}(j) > k, \end{cases}$$

and the proof is complete. □

Of course, if we substitute each x_j for $x_{(j)}$ in (2) we immediately obtain a symmetric spread measure. Such a class includes the *median absolute deviation*, given by the equation $\mathsf{MAD}(\mathbf{x}) = \mathsf{Med}\left(|\mathbf{x} - \mathsf{Med}(\mathbf{x})|\right)$.

What is more, a difference between sample quantiles $\mathsf{V}(\mathbf{x}) = \mathsf{Q}_{\alpha'}(\mathbf{x}) - \mathsf{Q}_{\alpha''}(\mathbf{x})$ for any $\alpha' \geq \alpha''$ is also included here. This is because if $\mathsf{V}(\mathbf{x}) = \gamma' x_{(k')} + (1 - \gamma')x_{(k'+1)} - \gamma'' x_{(k'')} - (1-\gamma'')x_{(k''+1)}$ for some $\gamma', \gamma'', k', k'', k'' < k'$, then $\mathsf{V}(\mathbf{x})$ may be expressed as $\gamma' \left|x_{(k')} - \mathsf{Q}_{\alpha''}(\mathbf{x})\right| + (1-\gamma') \left|x_{(k'+1)} - \mathsf{Q}_{\alpha''}(\mathbf{x})\right|$. Thus, by Proposition 14, it is a symmetric spread measure. Among the best known examples of such functions we find the *interquartile range* given by $\mathsf{IQR}(\mathbf{x}) = \mathsf{Q}_{0.75}(\mathbf{x}) - \mathsf{Q}_{0.25}(\mathbf{x})$ and of course the already mentioned *sample range*.

Proposition 14 indicates some new, possibly interesting spread measures like $\sum_{i=1}^{n}(x_i - \mathsf{Med}(\mathbf{x}))^2$ (if we would like to use it in statistics, then of course we should normalize its value somehow) or e.g. a non-necessarily symmetric function $\sum_{i=1}^{n} w_i\, |x_i - \mathsf{Med}(\mathbf{x})|$, which can be of potential interest in DM problems.

Additionally, note that as $|sx_i - \mathsf{Q}_\alpha(s\mathbf{x})| = s|x_i - \mathsf{Q}_\alpha(\mathbf{x})|$, V is homogeneous of degree 1 if and only if $\tilde{\mathsf{A}}$ fulfills (v4). For example, in the class of quasi-arithmetic means $\tilde{\mathsf{A}}(\mathbf{x}) = f^{-1}(\frac{1}{n}\sum_i f(x_i))$ only the geometric mean $(\prod_i x_i)^{1/n}$ and root-mean-power $(\frac{1}{n}\sum_i x_i^p)^{1/p}$ meets (v4), see (Grabisch et al., 2009, Theorem 4.15 (ii)).

5.3. Weighted average distance to a weighted average

Note that the function $\tilde{\mathsf{A}}(\mathbf{x}) = \left(\sum_{i=1}^{n} \frac{1}{n}|x_i|^p\right)^{1/p}$ is symmetric, fulfills (a2), and is nondecreasing on $\mathbb{I}^n$. In this vector space it coincides with the L_p-norm $\|\cdot\|_p$. Proposition 14 implies that, given the norm-generated metric d_p, it holds that $\mathsf{d}_p(\mathbf{x}, n{*}\mathsf{Q}_\alpha(\mathbf{x}))$ is a spread measure. Such a function has an intuitive interpretation: it denotes a distance to some sample quantile.

In this section we will study functions which generalize spread measures of this type. First of all, we will substitute Q_α for any weighted arithmetic mean (in the nonsymmetric case) or an OWA operator. Secondly, we will consider a weighted distance. We should note that, however, not all such functions fulfill (v1). For example, if $\mathsf{F}(\mathbf{x}) = \sum_{i=1}^{5} 0.2(x_i - (0.5x_1 + 0.5x_5))^2$, then $\mathsf{F}(0, 0.9, 0.9, 0.9, 0.9) = 0.2025$ and $\mathsf{F}(0, 0.9, 0.9, 0.9, 1) = 0.196$.

Let $\mathbf{w} \in [0,1]^n$ be a weighting vector, i.e. such that $\sum_{i=1}^{n} w_i = 1$. Recall that the *weighted arithmetic mean* is given by $\mathsf{WAM}_\mathbf{w}(\mathbf{x}) = \sum_{i=1}^{n} w_i\, x_i$. It is an idempotent aggregation function ($(\forall c \in \mathbb{I})\ \mathsf{WAM}_\mathbf{w}(n * c) = c$), for which for any $\mathbf{x} \in \mathbb{I}^n$ and s, t with $s\mathbf{x} + t \in \mathbb{I}^n$ it holds $\mathsf{WAM}_\mathbf{w}(s\mathbf{x} + t) = s\mathsf{WAM}_\mathbf{w}(\mathbf{x}) + t$.

If we substitute x_i for $x_{(i)}$ in the definition of $\mathsf{WAM}_\mathbf{w}$, then we obtain a symmetric aggregation function $\mathsf{OWA}_\mathbf{w}$ called *ordered weighted averaging operator*, see (Yager, 1988).

Definition 15. Given $p \geq 1$ and weighting vectors $\mathbf{v}, \mathbf{w}$, the $\mathsf{WD}_p\mathsf{WAM}_{\mathbf{v},\mathbf{w}}$ operator for $p < \infty$ is defined as

$$\mathsf{WD}_p\mathsf{WAM}_{\mathbf{v},\mathbf{w}}(\mathbf{x}) = \left(\sum_{i=1}^{n} v_i \left|x_i - \mathsf{WAM}_{\mathbf{w}}(\mathbf{x})\right|^p\right)^{1/p},$$

and for $p = \infty$ as:

$$\mathsf{WD}_p\mathsf{WAM}_{\mathbf{v},\mathbf{w}}(\mathbf{x}) = \max_{i=1,\dots,n} v_i \left|x_i - \mathsf{WAM}_{\mathbf{w}}(\mathbf{x})\right|.$$

We immediately have that $\mathsf{WD}_p\mathsf{WAM}$ operators fulfill the following properties.

Proposition 16. *For any $p \geq 1$, and all weighting vectors $\mathbf{v}$, $\mathbf{w}$, $\mathsf{WD}_p\mathsf{WAM}_{\mathbf{v},\mathbf{w}}$ fulfills (v0), (v2), (v3'), and (v4).*

The elementary proof is omitted.

In many practical cases we will often be interested in the $p \in \{1, 2, \infty\}$ case with the setting $v_i = w_i$.

Moreover $\mathsf{WD}_p\mathsf{OWA}_{\mathbf{v},\mathbf{w}}$ may be defined as symmetrized version of $\mathsf{WD}_p\mathsf{WAM}_{\mathbf{v},\mathbf{w}}$: in order to fulfill (v6) we substitute x_i for $x_{(i)}$; of course $v_i = w_i = \frac{1}{n}$ gives $\mathsf{WD}_p\mathsf{OWA}_{\mathbf{v},\mathbf{w}}(\mathbf{x}) = \mathsf{WD}_p\mathsf{WAM}_{\mathbf{v},\mathbf{w}}(\mathbf{x})$. Note that if $\mathsf{OWA}_{\mathbf{w}}$ denotes a sample quantile, then we may make use of Proposition 14.

In the following subsections we will examine when (v1) is fulfilled for $p = 1, 2, \infty$. Before that we shall provide the following auxiliary result. Fix $\sigma \in \mathfrak{S}_{[n]}$ and take any $\mathbf{x} \in \mathbb{I}^n_\sigma$. Assuming that $\boldsymbol{\delta} = \mathsf{diff}(\mathbf{x})$, we have:

$$\begin{aligned}
\mathsf{WAM}_{\mathbf{w}}|_\sigma(\mathbf{x}) &= \sum_{i=1}^{n} w_{\sigma(i)} \left(x_{\sigma(1)} + \sum_{j=1}^{i-1} \delta_j\right) = x_{\sigma(1)} + \sum_{i=2}^{n} \left(w_{\sigma(i)} \sum_{j=1}^{i-1} \delta_j\right) \\
&= x_{\sigma(1)} + \sum_{i=1}^{n-1} \left(\delta_i \sum_{j=i+1}^{n} w_{\sigma(j)}\right) = x_{\sigma(1)} + \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma, \qquad (3)
\end{aligned}$$

with $\zeta_i^\sigma = \sum_{j=i+1}^{n} w_{\sigma(j)}$, $i \in [n-1]$. Note that $\boldsymbol{\zeta}^\sigma \in [0,1]^{n-1}$ is nonincreasing and $1 - \zeta_i^\sigma = \sum_{j=1}^{i} w_{\sigma(j)}$.

5.3.1. $\mathsf{WD}_2\mathsf{WAM}$ *operators*

Let us begin with examining the $p = 2$ case. We have already noted that not all the functions of this kind fulfill (v1). In fact, we have what follows.

Proposition 17. *Fix* $\sigma \in \mathfrak{S}_{[n]}$*, weighting vectors* $\mathbf{v}, \mathbf{w}$*, and let* $\mathsf{V}(\mathbf{x}) = \sum_{i=1}^n v_i \left(x_i - \sum_{j=1}^n w_j x_j\right)^2$*. Then* $\mathsf{V}|_\sigma$ *fulfills (v1) if and only if* $\mathbf{v}$ *and* $\mathbf{w}$ *are such that for all* $1 \le a < b \le n-1$ *it holds:*

$$\left(\sum_{i=b+1}^n v_{\sigma(i)}\right)\left(\sum_{i=1}^a w_{\sigma(i)}\right) + \left(\sum_{i=1}^a v_{\sigma(i)}\right)\left(\sum_{i=b+1}^n w_{\sigma(i)}\right) \ge \left(\sum_{i=1}^a w_{\sigma(i)}\right)\left(\sum_{i=b+1}^n w_{\sigma(i)}\right).$$

Proof. Let $\mathbf{x} \in \mathbb{I}_\sigma^n$ and assume that $\boldsymbol{\delta} := \mathsf{diff}(\mathbf{x})$. By Eq. (3) for $\zeta_i^\sigma = \sum_{j=i+1}^n w_{\sigma(j)}$, $i \in [n-1]$, we have:

$$\mathsf{V}|_\sigma(\mathbf{x}) = \mathsf{V}|_\sigma(\boldsymbol{\delta}) = v_{\sigma(1)} \left(\sum_{i=1}^{n-1} \delta_i\, \zeta_i^\sigma\right)^2 + \sum_{i=1}^{n-1} v_{\sigma(i+1)} \left(\sum_{j=1}^{i} \delta_j - \sum_{j=1}^{n-1} \delta_j\, \zeta_j^\sigma\right)^2.$$

Temporarily, we may of course treat $\mathsf{V}|_\sigma$ as a function of $\boldsymbol{\delta}$. We have to determine for which $\mathbf{v}, \mathbf{w}$ the function $\mathsf{V}|_\sigma$ is nondecreasing w.r.t. each element in $\boldsymbol{\delta}$. As $\mathsf{V}|_\sigma$ is differentiable, we may examine all its partial derivatives. Assuming that $\xi_i^\sigma = \sum_{j=i+1}^n v_{\sigma(j)}$, $i \in [n-1]$, for any $a \in [n-1]$ it holds:

$$\begin{aligned}
\frac{1}{2}\frac{\partial}{\partial \delta_a} \mathsf{V}|_\sigma(\boldsymbol{\delta}) &= v_{\sigma(1)}\, \zeta_a^\sigma \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma + \sum_{i=1}^{n-1} v_{\sigma(i+1)} \left(\mathrm{Ind}(a \le i) - \zeta_a^\sigma\right)\left(\sum_{j=1}^{i} \delta_j - \sum_{j=1}^{n-1} \delta_j \zeta_j^\sigma\right) \\
&= v_{\sigma(1)} \zeta_a^\sigma \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma + \zeta_a^\sigma \sum_{i=1}^{n-1} v_{\sigma(i+1)} \sum_{j=1}^{n-1} \delta_j \zeta_j^\sigma - \sum_{i=a}^{n-1} v_{\sigma(i+1)} \sum_{j=1}^{n-1} \delta_j \zeta_j^\sigma \\
&\quad - \zeta_a^\sigma \sum_{i=1}^{n-1} v_{\sigma(i+1)} \sum_{j=1}^{i} \delta_j + \sum_{i=a}^{n-1} v_{\sigma(i+1)} \sum_{j=1}^{i} \delta_j \\
&= \left((\zeta_a^\sigma - \xi_a^\sigma) \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma\right) + \left(\sum_{i=1}^{n-1} \left((\xi_a^\sigma \wedge \xi_i^\sigma) - \xi_i^\sigma \zeta_a^\sigma\right) \delta_i\right) \\
&= \sum_{i=1}^{n-1} \delta_i \left(\zeta_i^\sigma \zeta_a^\sigma - \zeta_i^\sigma \xi_a^\sigma + \xi_a^\sigma \wedge \xi_i^\sigma - \xi_i^\sigma \zeta_a^\sigma\right) = (*).
\end{aligned}$$

We of course have to determine in which cases it holds $\partial \mathsf{V}|_\sigma / \partial \delta_a \ge 0$ for all a and $\boldsymbol{\delta}$. However, it is easily seen that it is necessary and sufficient to consider only the diff-vectors of the form $\boldsymbol{\delta} = (0, \ldots, 0, \delta_b, 0, \ldots, 0)$, i.e. with $\delta_b > 0$, $\delta_i = 0$, for all $b, i = 1, \ldots, n-1$ such that $b \neq i$.

Therefore, at this point of the proof we have already shown that $\mathsf{V}|_\sigma$ is a spread measure if and only if for all $a, b = 1, \ldots, n-1$ it holds $(*) \geq 0$. What is more, as $\boldsymbol{\xi}^\sigma$ is sorted nonincreasingly, we may note that the cases $(a, b) = (i, j)$ and $(a, b) = (j, i)$ are equivalent:

$$\begin{aligned} \text{(i) If } a > b, \text{ then} \quad (*) &= \xi_a^\sigma(1 - \zeta_b^\sigma) + \zeta_a^\sigma(\zeta_b^\sigma - \xi_b^\sigma); \\ \text{(ii) If } b < a, \text{ then} \quad (*) &= \xi_b^\sigma(1 - \zeta_a^\sigma) + \zeta_b^\sigma(\zeta_a^\sigma - \xi_a^\sigma). \end{aligned}$$

Additionally, for $a = b$ the condition is always true:

$$(*) = \zeta_a^{\sigma 2} - 2\xi_a^\sigma \zeta_a^\sigma + \xi_a^\sigma,$$

as this is an upward-open parabola with discriminant $\Delta(\xi_a^\sigma) = 4\,\xi_a^\sigma\,(\xi_a^\sigma - 1) \leq 0$.

Thus, we only have to check whether

$$\xi_b^\delta(1 - \zeta_a^\delta) + (\zeta_a^\delta - \xi_a^\delta)\zeta_b^\delta \geq 0$$

is true for $1 \leq a < b \leq n-1$.

Now let us rewrite the conditions so that they directly depend on $\mathbf{v}$ and $\mathbf{w}$ rather than $\boldsymbol{\xi}^\delta$ and $\boldsymbol{\zeta}^\delta$, respectively:

$$\begin{aligned} \xi_b^\sigma(1 - \zeta_a^\sigma) - \xi_a^\sigma \zeta_b^\sigma &\geq -\zeta_a^\sigma \zeta_b^\sigma \\ \xi_b^\sigma(1 - \zeta_a^\sigma) - \xi_a^\sigma \zeta_b^\sigma + \zeta_b^\sigma &\geq -\zeta_a^\sigma \zeta_b^\sigma + \zeta_b^\sigma \\ \xi_b^\sigma(1 - \zeta_a^\sigma) + (1 - \xi_a^\sigma)\zeta_b^\sigma &\geq (1 - \zeta_a^\sigma)\zeta_b^\sigma \\ \left(\sum_{i=b+1}^{n} v_{\sigma(i)}\right)\left(\sum_{i=1}^{a} w_{\sigma(i)}\right) + \left(\sum_{i=1}^{a} v_{\sigma(i)}\right)\left(\sum_{i=b+1}^{n} w_{\sigma(i)}\right) &\geq \left(\sum_{i=1}^{a} w_{\sigma(i)}\right)\left(\sum_{i=b+1}^{n} w_{\sigma(i)}\right). \end{aligned}$$

Thus, the proof is complete. □

Of course, in case of symmetric $\mathsf{WD_2OWA}_{\mathbf{v},\mathbf{w}}$ spread measures we should examine if the conditions mentioned in Proposition 17 hold only in the case $\sigma(i) = i$. For nonsymmetric ones, unfortunately, we should consider each possible permutation σ. However, it is easily seen that if $v_i = w_i$ for all $i \in [n]$, then $\mathsf{V}|_\sigma$ fulfills (v1) for every $\sigma \in \mathfrak{S}_{[n]}$. Thus, each $\mathsf{WD_2WAM}_{\mathbf{w},\mathbf{w}}$ is a spread measure. For example, after a proper transformation of $\mathsf{WD_2WAM}_{\mathbf{w},\mathbf{w}}$ in case of $v_i = w_i = \frac{1}{n}$ we again obtain the standard deviation and sample variance.

5.3.2. $\mathsf{WD}_1\mathsf{WAM}$ *operators*

Furthermore, let us consider the $p = 1$ case. We should note that not all $\mathsf{WD}_1\mathsf{WAM}$ operators fulfill (v1).

Example 4. Let $\mathsf{V}(\mathbf{x}) = \sum_{i=1}^n 0.2|x_i-(0.5x_1+0.5x_5)|$. We have $\mathsf{V}(0, 0.9, 0.9, 0.9, 0.9) = 0.45 > \mathsf{V}(0, 0.9, 0.9, 0.9, 1) = 0.44$.

The $\mathsf{WD}_1\mathsf{WAM}$ operators that fulfill (v1) are characterized in the following proposition.

Proposition 18. *Fix* $\sigma \in \mathfrak{S}_{[n]}$*, weighting vectors* $\mathbf{v}, \mathbf{w}$*, and let* $\mathsf{V}(\mathbf{x}) = \sum_{i=1}^n v_i \left|x_i - \sum_{j=1}^n w_i x_j\right|$*. The* $\mathsf{V}|_\sigma$ *fulfills (v1) if and only if* $\mathbf{v}, \mathbf{w}$ *are such that for all* $r, k \in [n-1]$ *it holds*

$$\sum_{i=1}^{r} w_{\sigma(i)} \left(\sum_{i=1}^{k} v_{\sigma(i)} - \sum_{i=k+1}^{n} v_{\sigma(i)} \right) \le \sum_{i=1}^{r} v_{\sigma(i)} \quad \textit{for} \quad r < k,$$
$$\sum_{i=r+1}^{n} w_{\sigma(i)} \left(\sum_{i=k+1}^{n} v_{\sigma(i)} - \sum_{i=1}^{k} v_{\sigma(i)} \right) \le \sum_{i=r+1}^{n} v_{\sigma(i)} \quad \textit{for} \quad r \ge k.$$

Proof. Let $\mathbf{x} \in \mathbb{I}_\sigma^n$ and $\boldsymbol{\delta} := \mathsf{diff}(\mathbf{x})$. Assuming that $\zeta_i^\sigma = \sum_{j=i+1}^n w_{\sigma(j)}$ and $\xi_i^\sigma = \sum_{j=i+1}^n v_{\sigma(j)}$, $i \in [n-1]$, let $k = \min\{k \in [n-1] : \sum_{i=1}^k \delta_i - \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma \ge 0\}$. By Eq. (3) we have:

$$\begin{aligned}
\mathsf{V}|_\sigma(\mathbf{x}) &= \mathsf{V}|_\sigma(\boldsymbol{\delta}) = v_{\sigma(1)} \left| \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma \right| + \sum_{i=1}^{n-1} v_{\sigma(i+1)} \left| \sum_{j=1}^{i} \delta_j - \sum_{j=1}^{n-1} \delta_j \zeta_j^\sigma \right| \\
&= v_{\sigma(1)} \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma + \sum_{i=1}^{k-1} v_{\sigma(i+1)} \left(\sum_{j=1}^{n-1} \delta_j \zeta_j^\sigma - \sum_{j=1}^{i} \delta_j \right) + \sum_{i=k}^{n-1} v_{\sigma(i+1)} \left(\sum_{j=1}^{i} \delta_j - \sum_{j=1}^{n-1} \delta_j \zeta_j^\sigma \right) \\
&= (1 - \xi_k^\sigma) \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma - \xi_k^\sigma \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma - \sum_{i=1}^{k-1} (\xi_i^\sigma - \xi_k^\sigma) \delta_i + \sum_{i=k}^{n-1} \xi_i^\sigma \delta_i + \sum_{i=1}^{k-1} \xi_k^\sigma \delta_i \\
&= \sum_{i=1}^{n-1} \Big((1 - 2\xi_k^\sigma) \zeta_i^\sigma + \mathrm{Ind}(i < k)(2\xi_k^\sigma - \xi_i^\sigma) + \mathrm{Ind}(i \ge k) \xi_i^\sigma \Big) \delta_i.
\end{aligned}$$

We shall find out for which $\boldsymbol{\zeta}^\sigma, \boldsymbol{\xi}^\sigma$ it holds $\mathsf{V}|_\sigma(\boldsymbol{\delta}') - \mathsf{V}|_\sigma(\boldsymbol{\delta}) \ge 0$ for any $r \in [n-1]$ and $\Delta > 0$, where $\boldsymbol{\delta}' = (\delta_1, \ldots, \delta_{r-1}, \delta_r + \Delta, \delta_{r+1}, \ldots, \delta_{n-1})$. Assuming that $k' = \min\{k' \in [n-1] : \sum_{i=1}^{k'} \delta_i' - \sum_{i=1}^{n-1} \delta_i' \zeta_i^\sigma \ge 0\}$, let us consider three cases.

(i) If $k' = k$, then we have:

$$\mathsf{V}|_\sigma(\boldsymbol{\delta}') - \mathsf{V}|_\sigma(\boldsymbol{\delta}) = ((1 - 2\xi_k^\sigma) \zeta_r^\sigma + \mathrm{Ind}(r < k)(2\xi_k^\sigma - \xi_r^\sigma) + \mathrm{Ind}(r \ge k) \xi_r^\sigma)\, \Delta.$$

First of all, we should note that for each $\boldsymbol{\zeta}^\sigma$ and k there always exists $\boldsymbol{\delta}$ such that for some $\boldsymbol{\delta}'$ it holds $k' = k$. Therefore, $\mathsf{V}|_\sigma$ in order to fulfill (v1) must necessarily be generated by $\boldsymbol{\zeta}^\sigma$ and $\boldsymbol{\xi}^\sigma$ such that:

$$\begin{aligned}(1-2\xi^\sigma_k)\zeta^\sigma_r + 2\xi^\sigma_k - \xi^\sigma_r \geq 0 \quad \text{if} \quad r < k,\\ (1-2\xi^\sigma_k)\zeta^\sigma_r + \xi^\sigma_r \geq 0 \quad \text{if} \quad r \geq k.\end{aligned}$$

for all $r, k \in [n-1]$. The condition for $r < k$ is equivalent to:

$$\begin{aligned}(1-2\xi^\sigma_k)\zeta^\sigma_r + 2\xi^\sigma_k - \xi^\sigma_r &\geq 0\\ (1-2\xi^\sigma_k)\zeta^\sigma_r - (1-2\xi^\sigma_k) + (1-\xi^\sigma_r) &\geq 0\\ (1-\zeta^\sigma_r)(1-2\xi^\sigma_k) &\leq (1-\xi^\sigma_r)\\ \sum_{i=1}^{r} w_{\sigma(i)}\left(\sum_{i=1}^{k} v_{\sigma(i)} - \sum_{i=k+1}^{n} v_{\sigma(i)}\right) &\leq \sum_{i=1}^{r} v_{\sigma(i)},\end{aligned}$$

and the second condition to:

$$\begin{aligned}(1-2\xi^\sigma_k)\zeta^\sigma_r + \xi^\sigma_r &\geq 0\\ \sum_{i=r+1}^{n} w_{\sigma(i)}\left(\sum_{i=k+1}^{n} v_{\sigma(i)} - \sum_{i=1}^{k} v_{\sigma(i)}\right) &\leq \sum_{i=r+1}^{n} v_{\sigma(i)}.\end{aligned}$$

(ii) Now we shall examine the case $k' > k$. Let us assume that conditions derived in case (i) hold. Note that we have $\mathsf{WAM}_{\mathbf{w}}(\mathbf{x}') - \mathsf{WAM}_{\mathbf{w}}(\mathbf{x}) = \zeta^\sigma_r\Delta$, where $x'_i = x_i + \mathrm{Ind}(i \geq r)\Delta$. By the definition of k' and k it holds $x_{\sigma(k')} - \mathsf{WAM}_{\mathbf{w}}(\mathbf{x}) \geq 0$ and $x_{\sigma(k')} + \mathrm{Ind}(r < k')\Delta - \mathsf{WAM}_{\mathbf{w}}(\mathbf{x}) - \zeta^\sigma_r\Delta < 0$. Thus, it must necessarily hold that $r \geq k'$ and $-\zeta^\sigma_r\Delta < \mathsf{WAM}_{\mathbf{w}}(\mathbf{x}) - x_{\sigma(k')} \leq 0$.

$$\begin{aligned}&\mathsf{V}|_\sigma(\boldsymbol{\delta}') - \mathsf{V}|_\sigma(\boldsymbol{\delta}) - \left(\mathrm{Ind}(r < k')(2\xi^\sigma_{k'} - \xi^\sigma_r) + \mathrm{Ind}(r \geq k')\xi^\sigma_r\right)\Delta\\ =\;& (1-2\xi^\sigma_{k'})\zeta^\sigma_r\Delta + (2\xi^\sigma_k - 2\xi^\sigma_{k'})\sum_{i=1}^{n-1}\zeta^\sigma_i\delta_i - \sum_{i=1}^{k-1}(2\xi^\sigma_k - 2\xi^\sigma_{k'})\delta_i - \sum_{i=k}^{k'-1}(2\xi^\sigma_i - 2\xi^\sigma_{k'})\delta_i\\ \geq\;& (1-2\xi^\sigma_{k'})\zeta^\sigma_r\Delta + 2(\xi^\sigma_k - \xi^\sigma_{k'})\left(\sum_{i=1}^{n-1}\delta_i\zeta^\sigma_i - \sum_{i=\alpha}^{k'-1}\delta_i\right)\\ =\;& (1-2\xi^\sigma_{k'})\zeta^\sigma_r\Delta + 2(\xi^\sigma_k - \xi^\sigma_{k'})(\mathsf{WAM}_{\mathbf{w}}(\mathbf{x}) - x_{\sigma(k')})\\ \geq\;& (1-2\xi^\sigma_{k'})\zeta^\sigma_r\Delta - 2(\xi^\sigma_k - \xi^\sigma_{k'})\zeta^\sigma_r\Delta = (1-2\xi^\sigma_k)\zeta^\sigma_r\Delta.\end{aligned}$$

As $r \geq k'$, it holds $\mathsf{V}|_\sigma(\boldsymbol{\delta}') - \mathsf{V}|_\sigma(\boldsymbol{\delta}) \geq (1-2\xi^\sigma_k)\zeta^\sigma_r\Delta + \xi^\sigma_r\Delta \geq 0$. We see that the necessary conditions derived in case (i) are sufficient in case (ii).

(iii) In the third case we assume that $k' < k$. Here we surely have $r \le k'$, as $\sum_{i=1}^{k'} \delta_i - \sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i < 0$ and $\sum_{i=1}^{k'} \delta_i + \mathrm{Ind}(r \le k')\Delta - \sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i - \zeta_r^\sigma \Delta \ge 0$.

$$
\begin{aligned}
& \mathsf{V}|_\sigma(\boldsymbol{\delta}') - \mathsf{V}|_\sigma(\boldsymbol{\delta}) - \left(\mathrm{Ind}(r < k')(2\xi_{k'}^\sigma - \xi_r^\sigma) + \mathrm{Ind}(r \ge k')\xi_r^\sigma\right)\Delta \\
= \;& (1 - 2\xi_{k'}^\sigma)\zeta_r^\sigma \Delta + (2\xi_k^\sigma - 2\xi_{k'}^\sigma) \sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i - \sum_{i=1}^{k'-1} (2\xi_k^\sigma - 2\xi_{k'}^\sigma)\delta_i - \sum_{i=k'}^{k-1} (2\xi_k^\sigma - 2\xi_i^\sigma)\delta_i \\
= \;& (1 - 2\xi_{k'}^\sigma)\zeta_r^\sigma \Delta + (2\xi_{k'}^\sigma - 2\xi_k^\sigma) \left(\sum_{i=1}^{k-1} \delta_i - \sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i \right) + \sum_{i=k'}^{k-1} (2\xi_{k'}^\sigma - 2\xi_i^\sigma)\delta_i \\
\ge \;& (1 - 2\xi_{k'}^\sigma)\zeta_r^\sigma \Delta + (2\xi_{k'}^\sigma - 2\xi_k^\sigma)(\zeta_r^\sigma - 1)\Delta + (2\xi_{k'}^\sigma - 2\xi_k^\sigma) \left(\sum_{i=1}^{k-1} \delta_i + \Delta - \sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i - \zeta_r^\sigma \right) \\
\ge \;& (1 - 2\xi_k^\sigma)\zeta_r^\sigma \Delta - (2\xi_{k'}^\sigma - 2\xi_k^\sigma)\Delta.
\end{aligned}
$$

As $r \le k'$ and for $r = k'$ it trivially holds $\xi_r^\sigma = 2\xi_{k'}^\sigma - \xi_r^\sigma$, we have:

$$
\begin{aligned}
\mathsf{V}|_\sigma(\boldsymbol{\delta}') - \mathsf{V}|_\sigma(\boldsymbol{\delta}) &\ge (1 - 2\xi_k^\sigma)\zeta_r^\sigma \Delta - (2\xi_{k'}^\sigma - 2\xi_k^\sigma)\Delta + (2\xi_{k'}^\sigma - \xi_r^\sigma)\Delta \\
&= (1 - 2\xi_k^\sigma)\zeta_r^\sigma \Delta + (2\xi_k^\sigma - \xi_r^\sigma)\Delta \ge 0.
\end{aligned}
$$

Thus, our necessary conditions are sufficient in all cases. □

Please be noticed that if $v_i = w_i$ for all $i \in [n]$, then the conditions in the above theorem are fulfilled for any σ. Thus, for all $\mathbf{w}$ the $\mathsf{WD_1WAM}_{\mathbf{w},\mathbf{w}}$ operator is a spread measure. For example, with this result we may imply that Fisher's (1922) mean error, $\mathsf{ME}(\mathbf{x}) = \frac{1}{n}\sqrt{\frac{\pi}{2}} \sum_{i=1}^n |x_i - \bar{x}|$ is a spread measure.

5.3.3. $\mathsf{WD}_\infty\mathsf{WAM}$ *operators*

In the $p = \infty$ case, $\mathbf{v} = \mathbf{w}$ does not imply that (v1) hold.

Example 5. Let $\mathbf{w} = \mathbf{v} = (0.2, 0, 0.5, 0, 0.05, 0.25)$, $\mathbf{x} = (0.12, 0.19, 0.77, 0.39, 0, 0.9)$, and $\mathbf{x}' = (0.22, 0.29, 0.87, 0.49, 0, 1)$. We have $\mathbf{x} \preccurlyeq_n \mathbf{x}'$, but $\mathsf{WD}_\infty\mathsf{WAM}(\mathbf{x})_{\mathbf{v},\mathbf{w}} \simeq 0.1028 > \mathsf{WD}_\infty\mathsf{WAM}(\mathbf{x}')_{\mathbf{v},\mathbf{w}} \simeq 0.1018$.

As here diverse values in $\mathbf{v}$ cause some interpretation difficulties, we will restrict ourselves to the case $v_i = 1/n$.

Proposition 19. *For any $\sigma \in \mathfrak{S}_{[n]}$ and weighting vector $\mathbf{w}$, let $\mathsf{V}(\mathbf{x}) = \max_{i=1,\dots,n} \left| x_i - \sum_{j=1}^n w_j x_j \right|$. Then $\mathsf{V}|_\sigma$ fulfills (v1).*

Proof. Let $\mathbf{x} \in \mathbb{I}^n_\sigma$ and $\boldsymbol{\delta} := \mathsf{diff}(\mathbf{x})$. Recalling how we have rewritten the definition of $\mathsf{WAM}_\mathbf{w}$ in Eq. (3) we have:

$$\mathsf{V}|_\sigma(\mathbf{x}) = \mathsf{V}|_\sigma(\boldsymbol{\delta}) = \sum_{i=1}^{n-1} \delta_i \zeta_i^\sigma \vee \bigvee_{i=1}^{n-1} \left| \sum_{j=1}^{i} \delta_j - \sum_{j=1}^{n-1} \delta_j \zeta_j^\sigma \right| = \sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i \vee \sum_{i=1}^{n-1} (1 - \zeta_i^\sigma) \delta_i.$$

We see that $\mathsf{V}|_\sigma$ is a nondecreasing function of elements in $\boldsymbol{\delta}$, thus it fulfills (v1). □

In other words, this result implies that if $\boldsymbol{\eta} = (n * \frac{1}{n})$, then $\mathsf{WD}_\infty\mathsf{WAM}_{\boldsymbol{\eta},\mathbf{w}}$ and $\mathsf{WD}_\infty\mathsf{OWA}_{\boldsymbol{\eta},\mathbf{w}}$ are spread measures.

6. Discussion

In this paper we have introduced a new family of functions that aim to synthesize numeric vectors: measures of (absolute) data spread, variability, or scatter. Although particular elements of this class have been studied in the literature for quite a long time – especially in probability and statistics – the presented results introduce them to the domain of aggregation theory and decision making. Thus, together with aggregation (averaging) functions and (fuzzy) logical connectives, spread measures are another examples of methods to reveal new "dimensions" of data being synthesized.

Of course, much still has to be done to understand the behavior of such measures better. Special attention should be paid to the study of the relation between spread measures for different vector arities (extended spread measures, i.e. functions in $\bigcup_{n=1}^\infty \mathbb{I}^n$, cf. Beliakov and James, 2013; Calvo and Mayor, 1999; Gagolewski and Grzegorzewski, 2011), possibly in some connection with multi-argument distances (Martin and Mayor, 2011).

Also, it is known that averaging aggregating functions may be obtained by minimizing some penalty P, which can be interpreted as a measure of consensus between the input vector and output value, see (Calvo and Beliakov, 2010; Mesiar, 2007), i.e. $\mathsf{A}(x) = \arg\min_y P(\mathbf{x}, y)$. Whether a similar nice interpretation may be given for spread measures (of course, not when we represent $\mathsf{V}(\mathbf{x})$ directly as $\tilde{\mathsf{A}}(\mathsf{diff}(\boldsymbol{\delta}))$, which leads to quite expectable outcomes) is definitely worth exploring.

Finally, we should note that exploratory data analysis also knows of some "normalized" measures of *relative* spread. For instance, the well known (unit-free) Gini coefficient, defined as $\mathsf{G}(\mathbf{x}) = \mathsf{MD}(\mathbf{x})/2\bar{\mathbf{x}}$ is definitely not a measure of absolute spread. This is because it does not even fulfill (v0): $\mathsf{G}(0, 2, 4) =$

$2/3$, and $\mathsf{G}(2,4,6) = 1/3$. Moreover, even though $(0,2,4) \preccurlyeq_n (0,3,5)$, we have $\mathsf{G}(0,3,5) = 5/8 < 2/3 = \mathsf{G}(0,2,4)$. A similar observation may be made about the so-called coefficient of variation, $\mathsf{CV}(\mathbf{x}) = \sqrt{\mathsf{Var}(\mathbf{x})}/\bar{\mathbf{x}}$. Both functions take into account the order of magnitude of the observations, and are ratio scale invariant (i.e. e.g. $\mathsf{CV}(s\mathbf{x}) = \mathsf{CV}(\mathbf{x})$) and continuous but not translation invariant.

More generally, the study of objects of the form $\mathsf{V}(\mathbf{x})/\mathsf{A}(\mathbf{x})$, where V is a spread measure, and A is an aggregation function, could be an interesting aim of future research.

Acknowledgments

The author would like to thank anonymous reviewers for valuable comments and suggestions on the manuscript and to Prof. Radko Mesiar and Anna Cena for stimulating discussion during the AGOP 2013 conference.

References

Aczel, A. (1996). *Complete Business Statistics*. Irvin.

Beliakov, G. and James, S. (2013). Stability of weighted penalty-based aggregation functions. *Fuzzy Sets and Systems*, 226(1):1–18.

Beliakov, G., Pradera, A., and Calvo, T. (2007). *Aggregation functions: A guide for practitioners*. Springer-Verlag.

Bernasconi, M., Choirat, C., and Seri, R. (2014). Empirical properties of group preference aggregation methods employed in AHP: Theory and evidence. *European Journal of Operational Research*, (232):584–592.

Calvo, T. and Beliakov, G. (2010). Aggregation functions based on penalties. *Fuzzy Sets and Systems*, 161:1420–1436.

Calvo, T. and Mayor, G. (1999). Remarks on two types of extended aggregation functions. *Tatra Mountains Mathematical Publications*, 16:235–253.

Calvo, T., Mayor, G., and Mesiar, R., editors (2002). *Aggregation operators. New trends and applications*, volume 97 of *Studies in Fuzziness and Soft Computing*. Physica-Verlag, New York.

Cramér, H. (1946). *Mathematical methods of statistics*. Princeton University Press, Princeton.

Fisher, R. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*, 222:309–368.

Gagolewski, M. and Grzegorzewski, P. (2011). Possibilistic analysis of arity-monotonic aggregation operators and its relation to bibliometric impact assessment of individuals. *International Journal of Approximate Reasoning*, 52(9):1312–1324.

Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2009). *Aggregation functions*. Cambridge University Press.

Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2011a). Aggregation functions: Construction methods, conjunctive, disjunctive and mixed classes. *Information Sciences*, 181:23–43.

Grabisch, M., Marichal, J.-L., Mesiar, R., and Pap, E. (2011b). Aggregation functions: Means. *Information Sciences*, 181:1–22.

Greco, S., Mesiar, R., and Rindone, F. (2014). Two new characterizations of universal integrals on the scale $[0, 1]$. *Information Sciences*. doi:10.1016/j.ins.2013.12.056.

Huang, Y.-S., Chang, W.-C., and Lin, Z.-L. (2013). Aggregation of utility-based individual preferences for group decision-making. *European Journal of Operational Research*, (229):462–469.

Hyndman, R. J. and Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician*, 50(4):361–365.

Kostal, L., Lansky, P., and Pokora, O. (2013). Measures of statistical dispersion based on Shannon and Fisher information concepts. *Information Sciences*, 235:214–223.

Martin, J. and Mayor, G. (2011). Multi-argument distances. *Fuzzy Sets and Systems*, 167:92–100.

Martín, J., Mayor, G., and Suñer, J. (2001). On dispersion measures. *Mathware & Soft Computing*, 8:227–237.

Mesiar, R. (2007). Fuzzy set approach to the utility, preference relations, and aggregation operators. *European Journal of Operational Research*, 176:414–422.

Ohki, M. and Murofushi, T. (2012). A ranking methodology using a new dispersion criterion on a group decision making. In *Proc. SCIS-ISIS 2012*, pages 1649–1653.

Pitman, E. (1939). The estimation of the location and scale parameters of a continuous population of any given form. *Biometrika*, 30:391–421.

Rothschild, M. and Stiglitz, J. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, 2(3):225–243.

Saaty, T. (1994). *Fundamentals of decision making and priority theory with the analytic hierarchy process*. RWS Publications, Pittsburgh.

Sanchez, D. and Trillas, E. (2012). Measures of fuzziness under different uses of fuzzy sets. In Greco, S. et al., editors, *Proc. IPMU 2012 (CCIS 298)*, pages 25–43. Springer-Verlag.

R Development Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Weber, S. (1984). Measures of fuzzy sets and measures of fuzziness. *Fuzzy Sets and Systems*, 13:247–271.

Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man, and Cybernetics*, 18(1):183–190.

Zeng, W. and Li, H. (2006). Inclusion measures, similarity measures, and the fuzziness of fuzzy sets and their relations. *International Journal of Intelligent Systems*, 21:639–653.