Normalized $\mathrm{WD}_p\mathrm{WAM}$ and $\mathrm{WD}_p\mathrm{OWA}$ spread measures

Marek Gagolewski[1,2]

[1] Systems Research Institute, Polish Academy of Sciences,
ul. Newelska 6, 01-447 Warsaw, Poland, gagolews@ibspan.waw.pl
[2] Faculty of Mathematics and Information Science, Warsaw University of Technology,
ul. Koszykowa 75, 00-662 Warsaw, Poland

Abstract

Aggregation theory often deals with measures of central tendency of quantitative data. As sometimes a different kind of information fusion is needed, an axiomatization of spread measures was introduced recently. In this contribution we explore the properties of $\mathsf{WD}_p\mathsf{WAM}$ and $\mathsf{WD}_p\mathsf{OWA}$ operators, which are defined as weighted L_p-distances to weighted arithmetic mean and OWA operators, respectively. In particular, we give forms of vectors that maximize such fusion functions and thus provide a way to normalize the output value so that the vector of maximal spread always leads to a fixed outcome, e.g., 1 if all the input elements are in $[0,1]$. This might be desirable when constructing measures of experts' opinions consistency or diversity in group decision making problems.

Keywords: data fusion, aggregation, spread, deviation, variance, OWA operators

1. Introduction

Let us recall the definition of an aggregation function, cf. [1, 2, 3, 4].

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) $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{A}(\mathbf{x}) = \sup \mathbb{I}$.

In particular, internal aggregation functions, i.e., such that $(\forall \mathbf{x})\ \mathsf{A}(\mathbf{x}) \in [\mathsf{Min}(\mathbf{x}), \mathsf{Max}(\mathbf{x})]$, are sometimes called **averaging functions**. The mentioned characteristic properties reflect somehow the concept of data synthesis: finding a value representative to the whole vector. Recently, the class of measures of absolute spread was introduced in [5]. A spread measure V may accompany an internal aggregation function A so that a numeric vector $\mathbf{x}$ is concisely described as $\mathsf{A}(\mathbf{x}) \pm \mathsf{V}(\mathbf{x})$. For different approaches to quantifying entropy or uncertainty of discrete probability mass functions see, e.g., [6, 7], cf. also the notion of fuzziness of a fuzzy set [8, 9, 10], multidiscances [11], a probability distribution's scale parameter estimates discussed by [12], or dissimilarity measures in case of $n = 2$ [13].

Recall that two vectors $\mathbf{x}, \mathbf{x}' \in \mathbb{I}^n$ are said to be **comonotonic** [2, Def. 2.123] if there exists a permutation σ of $[n] := \{1, 2, \dots, n\}$ such that

$$x_{\sigma(1)} \le \cdots \le x_{\sigma(n)} \quad \text{and} \quad x'_{\sigma(1)} \le \cdots \le x'_{\sigma(n)}.$$

Hence, the permutation σ orders the components of $\mathbf{x}$ and $\mathbf{x}'$ simultaneously. Another way to say that $\mathbf{x}$ and $\mathbf{x}'$ are comonotonic is that

$$(x_i - x_j)(x_i' - x_j') \ge 0 \text{ for every } i, j \in [n].$$

Interestingly, there exists an $O(n \log n)$-time algorithm for determining whether this property holds. Note that if all the elements in $\mathbf{x}$ are unique, then to determine if two vectors are comonotonic it is of course sufficient to take the (unique) ordering permutation σ of $\mathbf{x}$ and then verify if $(x'_{\sigma(1)}, \dots, x'_{\sigma(n)})$ is appropriately sorted. On the other hand, if there are tied observations in $\mathbf{x}$, we seek for subsequences of $\mathbf{x}$ such that $(x_{\sigma(i)}, x_{\sigma(i+1)}, \dots, x_{\sigma(i+k)})$ with $x_{\sigma(i)} = x_{\sigma(i+k)}$. Then we try to update σ so that it also sorts the corresponding observations in $\mathbf{x}'$. An exemplary implementation of such an algorithm is given in Fig. 1.

Let us introduce the following binary preordering relation $\preccurlyeq_n$ on $\mathbb{I}^n$, see [5]. 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 $\mathbf{x}$ and $\mathbf{x}'$ are comonotonic and for all $i, j \in [n]$ it holds:

$$|x_i - x_j| \le |x_i' - x_j'|.$$

It is easily seen that for all $s \ge 1$ and $\mathbf{x} \in \mathbb{I}^n$ such that $s\mathbf{x} = (sx_1, \dots, sx_n) \in \mathbb{I}^n$ we have $\mathbf{x} \preccurlyeq_n s\mathbf{x}$. What is more, for all $t \in \mathbb{R}$ for which $t + \mathbf{x} = (t + x_1, \dots, t + x_n) \in \mathbb{I}^n$ it holds $t + \mathbf{x} \preccurlyeq_n \mathbf{x} \preccurlyeq_n t + \mathbf{x}$. Additionally, for all $c \in \mathbb{I}$, $(n * c) = (c, c, \dots, c) \in \mathbb{I}^n$ is a minimal element of $(\mathbb{I}^n, \preccurlyeq_n)$.

The following class of fusion functions is an object of our main interest in this paper.

Definition 2. [5] 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 $\mathsf{V}(\mathbf{x}) \le \mathsf{V}(\mathbf{x}')$, (v2) for any $c \in \mathbb{I}$ it holds $\mathsf{V}(n * c) = 0$.

It may be shown that such descriptive statistics as sample variance, standard deviation, interquartile range, mean difference, median absolute deviation and functions like those analyzed, e.g., in [14] are spread measures.

Let us focus on a class of fusion functions that consists of objects defined as a weighted L_p-distance between a given vector and its weighted arithmetic mean. Given $p \in [1, \infty)$ and a weighting vector $\mathbf{w}$ ($(\forall i) w_i \geq 0$, $\sum_{i=1}^n w_i = 1$), the $\mathsf{WD}_p\mathsf{WAM}_\mathbf{w}$ operator is defined as

$$\mathsf{WD}_p\mathsf{WAM}_\mathbf{w}(\mathbf{x}) = \left(\sum_{i=1}^n w_i \big| x_i - \sum_{j=1}^n w_j x_j \big|^p \right)^{1/p}.$$

Moreover, if we substitute above each x_i with the i-th smallest value in $\mathbf{x}$, $x_{(i)}$, we get a $\mathsf{WD}_p\mathsf{OWA}_\mathbf{w}$ operator, cf. [15]:

$$\mathsf{WD}_p\mathsf{OWA}_\mathbf{w}(\mathbf{x}) = \left(\sum_{i=1}^n w_i \big| x_{(i)} - \sum_{j=1}^n w_j x_{(j)} \big|^p \right)^{1/p}.$$

Such a function is **symmetric**, i.e., it gives the same output value for all the permutations of elements in an input vector. Thus, it is particularly useful in statistics and data mining.

The paper is organized in the following manner. In the next Section we recall a characterization of spread measures and note that each such measures are closely related to aggregation functions acting on iterated differences between elements of an input vector. Next, we express the $\mathsf{WD}_2\mathsf{WAM}_\mathbf{w}$ and $\mathsf{WD}_1\mathsf{WAM}_\mathbf{w}$ operators in this very way. In Section 3 we introduce normalized spread measures so that the vectors of the greatest possible spread always result in the same output value. Finally, we conclude the paper in Section 4.

2. Spread measures as functions vectors' iterated differences

For any $\mathbf{x} \in \mathbb{I}^n$ 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. We have the following result, which becomes particularly appealing when we compare $(a1)$ and $(a2)$ with $(v1')$ and $(v2')$ below, respectively.

Theorem 3. *[5] $\mathsf{V} : \mathbb{I}^n \to [0, \infty]$ 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') $\inf_{\mathbf{x} \in \mathbb{I}^n} \mathsf{V}(\mathbf{x}) = 0$.

Let $\mathfrak{S}_{[n]}$ denote the set of all permutations of the set $[n]$. Given $\sigma \in \mathfrak{S}_{[n]}$, let $\mathbb{I}^n_\sigma = \{\mathbf{x} \in \mathbb{I}^n : x_{\sigma(1)} \leq \cdots \leq x_{\sigma(n)}\}$ denote the set of vectors in $\mathbb{I}^n$ such that σ is its ordering permutation and let $\mathcal{D}_\sigma = \{\mathsf{diff}(\mathbf{x}) : \mathbf{x} \in \mathbb{I}^n_\sigma\}$. We have $\mathcal{D}_\sigma = \{\boldsymbol{\delta} \in [0, b-a]^{n-1} : \sum_{i=1}^{n-1} \delta_i \leq b - a\} \subseteq [0, b-a]^{n-1}$.

What is more, denote by $\mathsf{V}|_\sigma$ the restriction of V to $\mathbb{I}^n_\sigma$, i.e., $\mathsf{V}|_\sigma(\mathbf{x}) = \mathsf{V}(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{I}^n_\sigma$. We see that 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, if $\mathbf{x} \in \mathbb{I}^n_\sigma$ and $\boldsymbol{\delta} = \mathsf{diff}(\mathbf{x})$, then $\mathsf{V}(\mathbf{x}) = \tilde{\mathsf{A}}_\sigma(\boldsymbol{\delta})$. In other words, we may define spread measures as $\preccurlyeq_n$ preserving mappings of x_i or, at the same time, nondecreasing functions ($\leq_{n-1}$-morphisms) of δ_i.

For instance, let us consider a fusion function defined as a difference between two sample quantiles

$$\begin{aligned} \mathsf{QD}(\mathbf{x}) &= \mathsf{Q}_{\alpha'}(\mathbf{x}) - \mathsf{Q}_{\alpha''}(\mathbf{x}) \\ &= (1 - \gamma')x_{(k')} + \gamma' x_{(k'+1)} \\ &- \gamma'' x_{(k'')} - (1 - \gamma'')x_{(k''+1)} \end{aligned}$$

for some $k' > k''$ and $\gamma', \gamma'' \in [0, 1]$ (see [16] for a review of definitions of quantiles in different statistical packages). After some simple transformations we get

$$\mathsf{QD}(\boldsymbol{\delta}) = \gamma'' \delta_{k''} + \sum_{i=k''+1}^{k'-1} \delta_i + \gamma' \delta_{k'}.$$

We see that it is a nondecreasing function of each δ_i and that for $\boldsymbol{\delta} = ((n-1) * 0)$ we have $\mathsf{QD}(\boldsymbol{\delta}) = 0$. Hence, it is a spread measure. Among particular instances of such functions we find the interquartile range, $\mathsf{IQR}(\mathbf{x}) = \mathsf{Q}_{0.75}(\mathbf{x}) - \mathsf{Q}_{0.25}(\mathbf{x})$, and range, $\mathsf{Range}(\boldsymbol{\delta}) = \sum_{i=1}^{n-1} \delta_i$.

In this section we are interested in redefining $\mathsf{WD}_p\mathsf{WAM}_\mathbf{w}$ and $\mathsf{WD}_p\mathsf{OWA}_\mathbf{w}$ as functions of $\boldsymbol{\delta}$ for the two most commonly used cases, $p = 1$ and $p = 2$.

Fix $\sigma \in \mathfrak{S}_n$ and let $\zeta^\sigma_i = \sum_{j=i+1}^n w_{\sigma(j)}$, $i \in [n-1]$. Note that $\boldsymbol{\zeta}^\sigma \in [0, 1]^{n-1}$ is nonincreasing and $1 - \zeta^\sigma_i = \sum_{j=1}^i w_{\sigma(j)}$. Additionally, we have that $\sum_{i=1}^n w_i x_i = x_{\sigma(1)} + \sum_{i=1}^{n-1} \delta_i \zeta^\sigma_i$.

2.1. $\mathsf{WD}_2\mathsf{WAM}$ as a function of $\boldsymbol{\delta}$

Let us start with the case $p = 2$.

Proposition 4. *For any weighting vector $\mathbf{w}$, $\mathsf{WD}_2\mathsf{WAM}_\mathbf{w}$ is a spread measure. Moreover, for any $\boldsymbol{\delta} \in \mathcal{D}_\sigma$ it holds:*

$$\mathsf{WD}_2\mathsf{WAM}_\zeta|_\sigma(\boldsymbol{\delta}) = \sqrt{\sum_{i=1}^{n-1} \sum_{k=1}^{n-1} \zeta^\sigma_{i \vee k}(1 - \zeta^\sigma_{i \wedge k}) \delta_i \delta_k}.$$

Proof. By [5, Proposition 17], we have:

$$\mathsf{WD}_2\mathsf{WAM}_\zeta|_\sigma(\boldsymbol{\delta})^2 = w_{\sigma(1)} \left(\sum_{i=1}^{n-1} \delta_i \, \zeta^\sigma_i \right)^2 + \sum_{i=1}^{n-1} w_{\sigma(i+1)} \left(\sum_{j=1}^i \delta_j - \sum_{j=1}^{n-1} \delta_j \, \zeta^\sigma_j \right)^2.$$

As for any $\boldsymbol{\delta} \in \mathcal{D}_\sigma$ it holds $\left(\sum_{j=1}^{i} \delta_j\right)^2 = \sum_{j=1}^{i} \delta_j^2 + 2\sum_{j=1}^{i-1}\sum_{k=j+1}^{i} \delta_j\delta_k = \sum_{j=1}^{i}\sum_{k=1}^{i} \delta_j\delta_k$, the function of our interest may be expressed as:

$$
\begin{aligned}
&\mathsf{WD_2WAM}_\zeta|_\sigma(\boldsymbol{\delta})^2 = \sum_{i=1}^{n} w_{\sigma(i)} \left(\sum_{i=1}^{n-1} \delta_i\, \zeta_i^\sigma\right)^2 \\
&-2\left(\sum_{i=1}^{n-1} \delta_i\zeta_i^\sigma\right)\left(\sum_{i=1}^{n-1} w_{\sigma(i+1)} \sum_{j=1}^{i} \delta_i\right) \\
&+\sum_{i=1}^{n-1} w_{\sigma(i+1)} \left(\sum_{j=1}^{i} \delta_i\right)^2 = \left(\sum_{i=1}^{n-1} \delta_i\, \zeta_i^\sigma\right)^2 \\
&-2\left(\sum_{i=1}^{n-1} \delta_i\zeta_i^\sigma\right)\left(\sum_{i=1}^{n-1} \delta_i\zeta_i^\sigma\right) + \sum_{i=1}^{n-1}\sum_{k=1}^{n-1} \delta_i\delta_k\zeta_{i\vee k}^\sigma \\
&=\sum_{i=1}^{n-1}\sum_{k=1}^{n-1} \delta_i\delta_k\zeta_i^\sigma\zeta_k^\sigma - 2\sum_{i=1}^{n-1}\sum_{k=1}^{n-1} \delta_i\delta_k\zeta_i^\sigma\zeta_k^\sigma \\
&+\sum_{i=1}^{n-1}\sum_{k=1}^{n-1} \delta_i\delta_k\zeta_{i\vee k}^\sigma = \sum_{i=1}^{n-1}\sum_{k=1}^{n-1} \delta_i\delta_k\left(\zeta_{i\vee k}^\sigma - \zeta_i^\sigma\zeta_k^\sigma\right) \\
&=\sum_{i=1}^{n-1}\sum_{k=1}^{n-1} \delta_i\delta_k\zeta_{i\vee k}^\sigma(1-\zeta_{i\wedge k}^\sigma).
\end{aligned}
$$

We see that $\mathsf{WD_2WAM}_\zeta|_\sigma(\boldsymbol{\delta})$ is a nondecreasing function of $\boldsymbol{\delta}$ and that for $\boldsymbol{\delta} = ((n-1) * 0)$ we have $\mathsf{WD_2WAM}_\zeta|_\sigma(\boldsymbol{\delta}) = 0$. Hence, it is a spread measure. $\square$

In particular, the **sample standard deviation** may be rewritten as:

$$
\mathsf{SD}(\mathbf{x}) = \sqrt{\sum_{i=1}^{n-1}\sum_{k=1}^{n-1} \frac{(i(n-k)) \wedge (k(n-i))}{n(n-1)}\delta_i\delta_k},
$$

which is a symmetric, property scaled (see Section 3) $\mathsf{WD_2OWA}$ spread measure.

Among examples of new $\mathsf{WD_2OWA}$ spread measures we may find, e.g.:

- **a k-trimmed standard deviation**, given by a weighting vector $\mathbf{w}$ such that $w_i = 0$ for $i \in \{1,\dots,k,n-k+1,\dots,n\}$, and $w_i = \frac{1}{n-2k}$ otherwise.
- **a k-Winsorized standard deviation**, given by a weighting vector $\mathbf{w}$ such that $w_i = 0$ for $i \in \{1,\dots,k,n-k+1,\dots,n\}$, $w_i = \frac{k+1}{n}$ for $i \in \{k+1, n-k\}$, and $w_i = \frac{1}{n}$ otherwise.

Both fusion functions are more robust in case of outliers imputed into an input data set. If we would like to use them as estimators of (population) standard deviation, they of course should be appropriately scaled in order to fulfill, e.g., asymptotic unbiasedness and/or consistency in particular statistical models. On the other hand, in a group decision making scenario, the 1-trimmed standard deviation might be used to measure the diversity of judges' scores in ski jumping competitions organized by the International Ski Federation, where each of 5 experts provide scores based on a jumper's balance, body position, and landing style. In such a case, one lowest and highest score is neglected.

2.2. $\mathsf{WD_1WAM}$ as a function of $\boldsymbol{\delta}$

Now we can turn our attention to the $p = 1$ case.

Proposition 5. *For any weighting vector* $\mathbf{w}$, $\mathsf{WD_1WAM}_\mathbf{w}$ *is a spread measure. Moreover, for any* $\boldsymbol{\delta} \in \mathcal{D}_\sigma$ *it holds:*

$$
\mathsf{WD_1WAM}_\zeta|_\sigma(\boldsymbol{\delta}) = 2\bigvee_{k=1}^{n-1}\sum_{i=1}^{n-1} \zeta_{i\vee k}^\sigma(1-\zeta_{i\wedge k}^\sigma)\delta_i.
$$

Proof. Assuming that $k = \min\{k' \in [n-1] : \sum_{i=1}^{k'} \delta_i \geq \sum_{i=1}^{n-1} \delta_i\zeta_i^\sigma\}$ by [5, Proposition 18], we have that:

$$
\begin{aligned}
\mathsf{WD_1WAM}_\zeta|_\sigma(\boldsymbol{\delta}) &= \sum_{i=1}^{n-1}\Big((1-2\zeta_k^\sigma)\zeta_i^\sigma \\
+ \quad &\mathrm{Ind}(i<k)(2\zeta_k^\sigma - \zeta_i^\sigma) + \mathrm{Ind}(i \geq k)\zeta_i^\sigma\Big)\delta_i.
\end{aligned}
$$

After few simple transformations we get that $\mathsf{WD_1WAM}_\zeta|_\sigma(\boldsymbol{\delta}) = \sum_{i=1}^{n-1} 2\zeta_{i\vee k}^\sigma(1-\zeta_{i\wedge k}^\sigma)\delta_i$. Assume $u(k') := \sum_{i=1}^{n-1} 2\zeta_{i\vee k'}^\sigma(1-\zeta_{i\wedge k'}^\sigma)\delta_i$. We shall show that $\bigvee_{k'=1}^{n-1} u(k') = u(k)$, i.e., that for any $e \in \mathbb{Z}$ such that $k+e \in [n-1]$ we have $u(k) \geq u(k+e)$.

1. Let $k' = k+e$, $e \geq 1$. Then:

$$
\begin{aligned}
&(u(k) - u(k+e))/2 \\
=&\left(\sum_{i=1}^{n-1} \zeta_i^\sigma\delta_i + \sum_{i=1}^{k-1}(\zeta_k^\sigma - \zeta_i^\sigma)\delta_i - \sum_{i=1}^{n-1} \zeta_i^\sigma\zeta_k^\sigma\delta_i\right) \\
-&\left(\sum_{i=1}^{n-1} \zeta_i^\sigma\delta_i + \sum_{i=1}^{k+e-1}(\zeta_{k+e}^\sigma - \zeta_i^\sigma)\delta_i - \sum_{i=1}^{n-1} \zeta_i^\sigma\zeta_{k+e}^\sigma\delta_i\right) \\
=&(\zeta_k^\sigma - \zeta_{k+e}^\sigma)\left(\sum_{i=1}^{k-1}\delta_i - \sum_{i=1}^{n-1}\zeta_i^\sigma\delta_i\right) \\
-&\sum_{i=k}^{k+e-1}(\zeta_{k+e}^\sigma - \zeta_i^\sigma)\delta_i \\
=&(\zeta_k^\sigma - \zeta_{k+e}^\sigma)\left(\sum_{i=1}^{k}\delta_i - \sum_{i=1}^{n-1}\zeta_i^\sigma\delta_i\right) \\
+&\sum_{i=k+1}^{k+e-1}(\zeta_i^\sigma - \zeta_{k+e}^\sigma)\delta_i \geq 0,
\end{aligned}
$$

as $\zeta_i^\sigma \geq \zeta_j^\sigma$ for $i \leq j$ and $\sum_{i=1}^{k}\delta_i - \sum_{i=1}^{n-1}\zeta_i^\sigma\delta_i \geq 0$ by the definition of k.

2. Let $k' = k - e$, $e \geq 1$. It holds:

$$\begin{aligned}
&(u(k) - u(k+e))/2 \\
=&(\zeta_k^\sigma - \zeta_{k-e}^\sigma)\left(\sum_{i=1}^{k-e-1} \delta_i - \sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i\right) \\
&+ \sum_{i=k-e}^{k-1} (\zeta_k^\sigma - \zeta_i^\sigma)\delta_i \\
=&(\zeta_{k-e}^\sigma - \zeta_k^\sigma)\left(\sum_{i=1}^{n-1} \zeta_i^\sigma \delta_i - \sum_{i=1}^{k-1} \delta_i\right) \\
&+ \sum_{i=k-e}^{k-1} (\zeta_{k-e}^\sigma - \zeta_i^\sigma)\delta_i \geq 0.
\end{aligned}$$

Additionally, we see that $\mathsf{WD_1WAM}_\zeta|_\sigma(\boldsymbol{\delta})$ is a nondecreasing function of $\boldsymbol{\delta}$ and that for $\boldsymbol{\delta} = ((n-1) * 0)$ we have $\mathsf{WD_1WAM}_\zeta|_\sigma(\boldsymbol{\delta}) = 0$. Therefore, it is a spread measure. □

For instance, Fisher's [17] **mean error**,

$$\begin{aligned}
\mathsf{ME}(\mathbf{x}) &= \frac{\sqrt{2\pi}}{2} \sum_{i=1}^{n} \frac{1}{n} \left| x_i - \frac{1}{n}\sum_{j=1}^{n} x_j \right| \\
&= \sqrt{2\pi} \bigvee_{k=1}^{n-1} \sum_{i=1}^{n-1} \frac{(n-(i \vee k))(i \wedge k)}{n^2} \delta_i
\end{aligned}$$

is a properly scaled (see below) $\mathsf{WD_1OWA}$ spread measure.

3. Normalized spread measures

A kind of "normalization" or scaling of a spread measure's output value is useful in many practical applications. For instance, when spread measures are utilized as point estimators of certain underlying probability distribution's characteristics, a proper transformation may lead to estimators fulfilling desirable properties. We know that, e.g., $\mathsf{V}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \sum_{j=1}^{n} x_j/n)^2$ is a biased estimator of the population variance and that $n\mathsf{V}(\mathbf{x})/(n-1)$ is free of such a systematic error. Of course, if V is a spread measure, then $\varphi \circ \mathsf{V}$ is also a spread measure for any nondecreasing $\varphi : [0, \infty] \to [0, \infty]$ such that $\varphi(0) = 0$.

In certain decision making problems, we may be interested in assuring that a vector of the greatest possible spread (as measured by a given V) leads to V's output value of $(b-a) = \sup \mathbb{I} - \inf \mathbb{I}$:

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

Each spread measure may be normalized in order to fulfill $(v3)$. It is because we have what follows.

Lemma 6. *Let V be a spread measure with $\sup_{\mathbf{x} \in \mathbb{I}^n} \mathsf{V}(\mathbf{x}) = u$. Then for each nondecreasing function $\varphi : [0, u] \to [0, (b-a)]$ such that $\varphi(0) = 0$, $\varphi(u) = b - a$, it holds that $\varphi \circ \mathsf{V}$ is a spread measure fulfilling (v3).*

A simple proof is omitted.

Clearly, each spread measure maximizes at some $\boldsymbol{\delta}$ such that $\sum_{i=1}^{n-1} \delta_i = b - a$, i.e., at $\mathbf{x}$ with $x_i = a$ and $x_j = b$ for some $i \neq j$. Thus, when determining the upper bound for a spread measure, we are faced with a constrained optimization problem.

For instance, a difference between two sample quantiles $\mathsf{QD}(\boldsymbol{\delta}) = \gamma'' \delta_{k''} + \sum_{i=k''+1}^{k'-1} \delta_i + \gamma' \delta_{k'}$ is maximized at $\boldsymbol{\delta}$ such that $\sum_{i=k''+1}^{k'-1} \delta_i = (b-a)$. On the other hand, a spread measure $\mathsf{V}(\boldsymbol{\delta}) = \bigwedge_{i=1}^{n-1} \delta_i$ is maximized if $(\forall i)\delta_i = (b-a)/(n-1)$ and $\mathsf{V}(\boldsymbol{\delta}) = \bigvee_{i=1}^{n-1} \delta_i$ reaches its maximum for $\boldsymbol{\delta}$ such that $\delta_i = (b-a)$ for some i and $\delta_j = 0$ for $j \neq i$, i.e., for $\mathbf{x} \in \{a, b\}^n$. This results may lead us to normalized versions of such spread measures that fulfill $(v3)$.

3.1. Normalized $\mathsf{WD_2WAM}$

We will show that $\sup_{\mathbf{x} \in \mathbb{I}^n} \mathsf{WD_2WAM}_{\mathbf{w}}(\mathbf{x}) \in [0, (b-a)/2]$, where the minimal possible value is obtained, e.g., for $\mathbf{w}$ such that for some i we have $w_i = 1$. On the other hand, we get the greatest value, e.g., for $\mathbf{w}$ such that $w_1 = w_n = 0.5$. Moreover, a vector with elements in $\{a, b\}^n$ (i.e., consisting of extreme values only) is of the greatest possible spread.

Theorem 7. *For any weighting vector $\mathbf{w}$ it holds $\arg\sup_{\mathbf{x} \in \mathbb{I}^n} \mathsf{WD_2WAM}_{\mathbf{w}}(\mathbf{x}) \in \{a, b\}^n$ and $\sup_{\mathbf{x} \in \mathbb{I}^n} \mathsf{WD_2WAM}_{\mathbf{w}}(\mathbf{x}) = (b-a)\sqrt{p(1-p)}$, where $p = \max_{A \subseteq [n], \sum_{i \in A} w_i \leq 0.5} \sum_{i \in A} w_i$.*

Proof. For simplicity, assume that $\mathsf{V}(\mathbf{x}) = \mathsf{WD_2WAM}_{\mathbf{w}}(\mathbf{x})^2$. Our optimization task may be written as:

Maximize $\mathsf{V}(\mathbf{x})$

subject to $a \leq x_i \leq b$, $i = 1, \ldots, n$.

By rewriting the constraints in terms of the Karush-Kuhn-Thucker (KKT) theorem we obtain $2n$ inequality constraints of the form:

$$\begin{aligned}
g_i(\mathbf{x}) = x_i - b &\leq 0, \\
g_{i+n}(\mathbf{x}) = a - x_i &\leq 0,
\end{aligned}$$

for $i = 1, \ldots, n$. Note that $\partial g_i(\mathbf{x})/\partial x_k = 1$ if $i = k$ and 0 otherwise and $\partial g_{n+i}(\mathbf{x})/\partial x_k = -1$ if $i = k$ and 0 otherwise. Moreover,

$$\frac{\partial}{\partial x_k} \mathsf{V}(\mathbf{x}) = 2w_k \left(x_k - \sum_{i=1}^{n} w_i x_i \right).$$

Hessian matrix $H(\mathsf{V})$, $h_{i,j} = \partial^2 \mathsf{V}(\mathbf{x})/\partial x_i \partial x_j$ is such that $h_{i,i} = 2w_i(1-w_i) \geq 0$ and $h_{i,j} = -2w_i w_j \leq 0$, $i \neq j$. We see that it is symmetric diagonally dominant as for any i we have $|h_{i,i}| = 2w_i(1-w_i) \geq \sum_{j \neq i} |h_{i,j}| = 2w_i \sum_{j \neq i} w_j = 2w_i(1-w_i)$. Thus, it is positive semi-definite, and we have that V is convex.

By the KKT theorem, if $\mathbf{x}^*$ is a local maximum of V satisfying the linear (affine) constraints

$g_i(\mathbf{x}^*) \le 0$, $g_{n+i}(\mathbf{x}^*) \le 0$, $i = 1,\dots,n$, then there exist $\mu_1,\dots,\mu_{2n} \ge 0$ such that:

$$\nabla \mathsf{V}(\mathbf{x}^*) = \sum_{i=1}^{n} \left(\mu_i \nabla g_i(\mathbf{x}^*) + \mu_{n+i}\nabla g_{n+i}(\mathbf{x}^*)\right)$$

and for $i = 1,\dots,n$

$$\mu_i g_i(\mathbf{x}^*) = 0, \quad \mu_{n+i} g_{n+i}(\mathbf{x}^*) = 0.$$

Thus,

$$\begin{cases} 2w_1\left(x_1^* - \sum_{i=1}^n w_i x_i^*\right) &= \mu_1 - \mu_{n+1} \\ 2w_2\left(x_2^* - \sum_{i=1}^n w_i x_i^*\right) &= \mu_2 - \mu_{n+2} \\ \vdots & \vdots \quad \vdots \\ 2w_n\left(x_n^* - \sum_{i=1}^n w_i x_i^*\right) &= \mu_n - \mu_{n+n} \end{cases}$$

with

$$\begin{cases} \mu_1(x_1^* - b) &= 0 \\ \vdots & \vdots \quad \vdots \\ \mu_n(x_n^* - b) &= 0 \\ \mu_{n+1}(a - x_i^*) &= 0 \\ \vdots & \vdots \quad \vdots \\ \mu_{n+n}(a - x_m^*) &= 0 \\ \mu_1 &\ge 0 \\ \vdots & \vdots \quad \vdots \\ \mu_n &\ge 0 \end{cases}$$

If $x_i^* = a$, then $\mu_i = 0$ and $\mu_{n+i} = \sum_{j=1}^n w_j x_j^* - a \ge 0$. If $x_i^* = b$, then $\mu_{n+i} = 0$ and $\mu_i = b - \sum_{j=1}^n w_j x_j^* \ge 0$. If $x_i^* \notin \{a, b\}$, then $\mu_i = \mu_{n+i} = 0$ and necessarily $x_i^* = \sum_{j=1}^n w_j x_j^*$. Thus, $\mathbf{x}^*$ maximizes V w.r.t. our constraints if it is such that $x_i^* \in \{a, b, \sum_{j=1}^n w_j x_j^*\}$.

As V is translation invariant, from now on – with no loss in generality – we may assume that $\mathbb{I} = [0,(b-a)]$. Take any index sets $A, B \subseteq [n]$, $A \cap B = \emptyset$. Let $\mathbf{x}$ be such that $x_i = 0$ for $i \in A$, $x_i = (b-a)$ for $i \in B$ and $x_i = \bar{x}$ otherwise, where

$$\begin{aligned} \bar{x} &= \sum_{i=1}^n w_i x_i \\ &= \sum_{i\in A} 0 w_i + \sum_{i\in B}(b-a)w_i + \sum_{i\in\overline{A\cup B}} \bar{x} w_i \\ &= (b-a)\frac{\sum_{i\in B} w_i}{\sum_{i\in A} w_i + \sum_{i\in B} w_i}. \end{aligned}$$

Moreover,

$$\begin{aligned} \mathsf{V}(\mathbf{x}) &= \sum_{i\in A} w_i \left((b-a)\frac{\sum_{i\in B} w_i}{\sum_{i\in A} w_i + \sum_{i\in B} w_i}\right)^2 \\ &+ \sum_{i\in B} w_i \left((b-a) - (b-a)\frac{\sum_{i\in B} w_i}{\sum_{i\in A} w_i + \sum_{i\in B} w_i}\right)^2 \\ &= \frac{(b-a)^2}{\left(\sum_{i\in A} w_i + \sum_{i\in B} w_i\right)^2} \\ &\times \left[\left(\sum_{i\in A} w_i\right)\left(\sum_{i\in B} w_i\right)^2 + \left(\sum_{i\in B} w_i\right)\left(\sum_{i\in A} w_i\right)^2\right] \\ &= \frac{(b-a)^2}{\left(\sum_{i\in A} w_i + \sum_{i\in B} w_i\right)^2} \\ &\times \left(\sum_{i\in A} w_i\right)\left(\sum_{i\in B} w_i\right)\left(\sum_{i\in B} w_i + \sum_{i\in A} w_i\right) \\ &= (b-a)^2 \frac{\left(\sum_{i\in A} w_i\right)\left(\sum_{i\in B} w_i\right)}{\sum_{i\in A} w_i + \sum_{i\in B} w_i}. \end{aligned}$$

It is easily seen that if A is fixed, then V maximizes for $B = \overline{A}$. Thus, $\mathbf{x}^* \in \{a, b\}^n$. In such a case we have $\mathsf{V}(\mathbf{x}^*) = (b-a)^2(\sum_{i\in A} w_i)(1 - \sum_{i\in A} w_i) \le (b-a)^2/4$, and the proof is complete. □

Thus, in order to compute $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{WD_2OWA}(\mathbf{x})$, we shall seek for $A \subseteq [n]$ that maximizes $f(A) = \sum_{i\in A} w_i$ subject to $f(A) \le 0.5$. This may be expressed as a binary programming task of finding $p = \max_{(b_1,\dots,b_n)\in\{0,1\}^n} \sum_{i=1}^n b_i w_i$ such that $\sum_{i=1}^n b_i w_i \le 0.5$.

On the other hand, from the proof of the above theorem it holds that in case of a $\mathsf{WD_2OWA}$ operator, one simply seeks for $p = \max_{k, \sum_{i\in[k]} w_i \le 0.5} \sum_{i\in[k]} w_i$, which can simply be computed in $O(n)$ time.

Basing on the above result, we may introduce the normalized $\mathsf{WD_2WAM_w}$ operator with $\mathbf{w}$ such that $(\forall i) w_i < 1$:

$$\mathsf{NWD_2WAM_w}(\mathbf{x}) = \frac{\sqrt{\sum_{i=1}^n w_i \left(x_i - \sum_{j=1}^n w_j x_j\right)^2}}{(b-a)\sqrt{p(1-p)}},$$

where $p = \max_{A\subseteq[n], \sum_{i\in A} w_i \le 0.5} \sum_{i\in A} w_i$. Also $\mathsf{NWD_2OWA_w}$ may be defined in a similar manner. Both of them of course fulfill $(v3)$. They also are **homogeneous of degree 1** (scale equivariant), see [2, Def. 2.86]), i.e., they meet:

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

and they are

(v5) continuous in each variable.

For instance, the sample variance is maximized, e.g., at $(\lfloor \frac{n}{2} \rfloor * b, \lceil \frac{n}{2} \rceil * a)$. Thus, its normalized version is given by:

$$\mathsf{NVar}(\mathbf{x}) = \eta\frac{2}{b-a}\sqrt{\sum_{i=1}^{n}\frac{1}{n}\left(x_i - \sum_{j=1}^{n}\frac{1}{n}x_j\right)^2},$$

where $\eta = 1$ for even n and $\eta = \sqrt{\frac{n^2}{(n-1)(n+1)}}$ otherwise.

3.2. Normalized $\mathsf{WD_1WAM}$

Let us show that $\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{WD_1WAM}_{\mathbf{w}}(\mathbf{x}) \in [0,(b-a)/2]$, and that a vector of the greatest possible spread is in $\{a,b\}^n$.

Theorem 8. *For any weighting vector* $\mathbf{w}$ *we have* $\arg\sup_{\mathbf{x}\in\mathbb{I}^n} \mathsf{WD_1WAM}_{\mathbf{w}}(\mathbf{x}) \in \{a,b\}^n$ *and* $\sup_{x\in\mathbb{I}^n} \mathsf{WD_1WAM}_{\mathbf{w}}(\mathbf{x}) = 2(b-a)p(1-p)$, *where* $p = \max_{A\subseteq[n],\sum_{i\in A} w_i\le 0.5}\sum_{i\in A} w_i$.

Proof. Fix $\sigma \in \mathfrak{S}_{[n]}$. Recall that we have $\mathsf{WD_1WAM}_{\boldsymbol{\zeta}}|_\sigma(\boldsymbol{\delta}) = 2\bigvee_{k=1}^{n-1}\sum_{i=1}^{n-1}\zeta^\sigma_{i\vee k}(1-\zeta^\sigma_{i\wedge k})\delta_i$. For fixed $k' \in [n-1]$ we have:

$$\begin{aligned}&\arg\max_{i\in[n-1]}\zeta^\sigma_{i\vee k'}(1-\zeta^\sigma_{i\wedge k'})\\ =&\arg\max_{i\in[n-1]}\left(\sum_{j=1}^{i\wedge k'} w_{\sigma(j)}\right)\left(\sum_{j=(i\vee k')+1}^{n} w_{\sigma(j)}\right) = k'.\end{aligned}$$

Note again that a maximal value of $\mathsf{WD_1WAM}_{\mathbf{w}}|_\sigma(\boldsymbol{\delta})$ is obtained for $\boldsymbol{\delta}$ such that $\sum_{i=1}^{n-1}\delta_i = (b-a)$. Thus, we have:

$$\sup_{\mathbf{x}\in\mathbb{I}^n}\mathsf{WD_1WAM}_{\boldsymbol{\zeta}}|_\sigma(\boldsymbol{\delta}) = 2\bigvee_{k=1}^{n-1}\zeta^\sigma_k(1-\zeta^\sigma_k)(b-a)$$

and we get that $\arg\sup_{\mathbf{x}\in\mathbb{I}^n}\mathsf{WD_1WAM}_{\mathbf{w}}|_\sigma(\mathbf{x}) \in \{a,b\}^n$.

Therefore, by considering each permutation $\sigma \in \mathfrak{S}_{[n]}$, we can conclude that it holds $\sup_{\mathbf{x}\in\mathbb{I}^n}\mathsf{WD_1WAM}_{\mathbf{w}}(\mathbf{x}) = 2(b-a)p(1-p)$, where $p = \max_{A\subseteq[n],\sum_{i\in A} w_i\le 0.5}\sum_{i\in A} w_i$. □

We may introduce a normalized version of $\mathsf{WD_1WAM}_{\mathbf{w}}$ for $\mathbf{w}$ such that $(\forall i) w_i < 1$ like:

$$\mathsf{NWD_1WAM}_{\mathbf{w}}(\mathbf{x}) = \frac{\sum_{i=1}^n w_i\left|x_i - \sum_{j=1}^n w_j x_j\right|}{2p(1-p)(b-a)},$$

where $p = \max_{A\subseteq[n],\sum_{i\in A} w_i\le 0.5}\sum_{i\in A} w_i$. Like before, also $\mathsf{NWD_1OWA}_{\mathbf{w}}$ operators may be defined.

Therefore, e.g., a scaled mean error that fulfills the $(v3)$ property may be defined as:

$$\mathsf{NME}(\mathbf{x}) = \eta\frac{2}{b-a}\sum_{i=1}^n\frac{1}{n}\left|x_i - \frac{1}{n}\sum_{j=1}^n x_j\right|,$$

where $\eta = 1$ for even n and $\eta = \frac{n^2}{(n-1)(n+1)}$ otherwise.

Note that all the spread measures studied in this subsection are homogeneous of degree 1.

```
#include <Rcpp.h>
#include <algorithm>
using namespace Rcpp;
// [[Rcpp::plugins("cpp11")]]

struct Comparer {
   const double* v;
   Comparer(const double* _v) { v = _v; }
   bool operator()(const int& i,
                   const int& j) {
      /* returns true if the first argument
        is less than (i.e., is ordered
        before) the second. */
      return (v[i] < v[j]);
   }
};

// [[Rcpp::export]]
bool is_comonotonic(NumericVector x,
                    NumericVector y) {
   int n = x.size();
   if (y.size() != n) stop("nx != ny");

   // recall that array elements in C++
   // are numbered from 0
   // let s = (0,1,...,n-1)
   std::vector<int> s(n);
   for (int i=0; i<n; ++i) s[i] = i;

   Comparer lt_x(REAL(x));
   std::sort(s.begin(), s.end(), lt_x);
   // now s is an ordering permutation of x

   Comparer lt_y(REAL(y));
   int i1 = 0;
   while (i1 < n) { /* now search for the
          longest subsequence consisting
          of equal x's */
      int i2 = i1+1;
      while (i2 < n && x[s[i1]] == x[s[i2]])
         ++i2;
      // sort the subsequence if necessary:
      if (i2-i1 > 1)
         std::sort(s.begin()+i1,
                   s.begin()+i2, lt_y);
      /* if y[s[i1-1]]>y[s[i1]] then x and y
        are not comonotonic: */
      if (i1 > 0 && y[s[i1-1]] > y[s[i1]])
         return false;
      i1 = i2;
   }

   /* as a by-product,
     (s[0]+1, s[1]+1, ..., s[n-1]+1) is a
     permutation that orders both x and y */
   return true;
}
```

Figure 1: A C++ (using Rcpp [19] classes) implementation of an $O(n\log n)$ algorithm to determine if two vectors of length n are comonotonic.

4. Conclusions

Measures of absolute spread are useful tools accompanying averaging aggregation functions in data analysis, data mining, and descriptive statistics. They may be used to describe a univariate data set concisely as $\mathsf{A}(\mathbf{x}) \pm \mathsf{V}(\mathbf{x})$. In such a case we most often rely on symmetric fusion functions defined on the space of vectors with elements in $\mathbb{I} = [-\infty,\infty]$.

In these types of applications, for example, some spread measures may also be used to determine the degree of outlyingness of data points: in [18], the sample median and median absolute deviation is utilized.

On the other hand, in decision making problems, where we restrict ourselves to the case of, e.g., $\mathbb{I} = [0, 1]$, nonsymmetric spread measures may be effectual. In this paper we introduced normalized versions of two important classes of such fusion functions so that their greatest possible value is equal to $\sup \mathbb{I} - \inf \mathbb{I}$. Such spread measures may provide an information on experts' opinions diversity.

An interesting direction for future research concerns the class of measures of *relative* spread, where the aggregation result is dependent on the order of magnitude of some averaging function A. This is the case of, e.g., the Gini coefficient, $\mathsf{G}(\mathbf{x}) = \mathsf{MD}(\mathbf{x})/2\bar{\mathbf{x}}$, or the coefficient of variation, $\mathsf{CV}(\mathbf{x}) = \sqrt{\mathsf{Var}(\mathbf{x})}/\bar{\mathbf{x}}$.

References

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

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

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

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

[5] M. Gagolewski. Spread measures and their relation to aggregation functions. *European Journal of Operational Research*, 241(2):469–477, 2015.

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

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

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

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

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

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

[12] E.J.G. Pitman. The estimation of the location and scale parameters of a continuous population of any given form. *Biometrika*, 30:391–421, 1939.

[13] H. Bustince, J. Fernandez, R. Mesiar, A. Pradera, and G. Beliakov. Restricted dissimilarity functions and penalty functions. In Sylvie Galichet et al., editors, *Proc. Eusflat/LFA 2011*, pages 79–85, 2011.

[14] P.J. Rousseeuw and C. Croux. Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.

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

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

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

[18] D.L. Donoho and M. Gasko. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *The Annals of Statistics*, 20(4):1803–1827, 1992.

[19] D. Eddelbuettel. *Seamless R and C++ Integration with Rcpp*. Springer, New York, 2013.