Information Sciences 300 (2015) 70–84

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

On the weak monotonicity of Gini means and other mixture functions

Gleb Beliakov [a,*], Tomasa Calvo [b], Tim Wilkin [a]

[a] School of Information Technology, Deakin University, 221 Burwood Hwy, Burwood 3125, Australia
[b] Department of Computer Sciences, Universidad de Alcalá, Alcala de Henares, Madrid, Spain

ARTICLE INFO

Article history:
Received 14 May 2014
Received in revised form 5 December 2014
Accepted 10 December 2014
Available online 29 December 2014

Keywords:
Aggregation function
Monotonicity
Mean
Penalty-based function
Gini mean
Mixture function

ABSTRACT

Weak monotonicity was recently proposed as a relaxation of the monotonicity condition for averaging aggregation, and weakly monotone functions were shown to have desirable properties when averaging data corrupted with outliers or noise. We extended the study of weakly monotone averages by analyzing their ϕ-transforms, and we established weak monotonicity of several classes of averaging functions, in particular Gini means and mixture operators. Mixture operators with Gaussian weighting functions were shown to be weakly monotone for a broad range of their parameters. This study assists in identifying averaging functions suitable for data analysis and image processing tasks in the presence of outliers.

1. Introduction

Growing interest in the field of aggregation functions [9,16,29] has led to the development of several new mathematical techniques, which have subsequently fostered the creation and analysis of new families of aggregation functions. One such technique is based on minimizing penalties that describe disagreement between the inputs [13,21], such that the output of an aggregation procedure is a value that is representative of the inputs in terms of the smallest penalty. It is well known that the arithmetic mean and the median are values that minimize the sum of squared and absolute deviations from the inputs respectively. Many other aggregation functions can be represented in terms of minima of particular penalty functions and all averaging functions can be represented as penalty based functions [13].

The class of penalty based functions is extensive and includes idempotent functions that are not monotone; for example the mode. In [10,14,32,34] a weaker condition than monotonicity was proposed for averaging functions, in which the value of the aggregate does not decrease when all the inputs are increased by the same value, but may decrease if a subset of the inputs increases. Two principal justifications for this weaker monotonicity were given: (a) many important, existing averages are not actually monotone, but rather are weakly monotone (for example the mode and robust estimators of location); and (b) weak monotonicity is very useful when calculating representative values of clusters of data in the presence of outliers. Indeed, cluster structure may change when only some inputs are increased (or decreased), but it does not change when all inputs are changed by the same value. It was shown that location estimators [27] and some useful classes of means, like

* Corresponding author.
E-mail addresses: gleb@deakin.edu.au (G. Beliakov), tomasa.calvo@uah.es (T. Calvo), tim.wilkin@deakin.edu.au (T. Wilkin).

http://dx.doi.org/10.1016/j.ins.2014.12.030
0020-0255/

Lehmer means, certain mixture functions, density-based means and the mode, are all weakly monotone [10,14,32,34]. The relation to the form of penalty functions for shift-invariance and hence weak monotonicity was established.

In this work we extend the results of [10,32] to encompass more general classes of means, in particular Gini means, as well as other mixture functions. Analysis of monotonicity of mixture and quasi-mixture functions has been performed recently in [19,24]. Mixture functions are a subclass of Bajraktarevic means, which generalize quasi-arithmetic means [2,3]. In mixture functions the inputs are averaged as in a weighted mean, but the weights depend on the inputs. Weights can thus be chosen so as to alternatively emphasize or de-emphasize the small or large inputs. When applied to averaging functions measuring the distance between the inputs, such means allow for de-emphasizing contributions from outliers. An equivalent class of functions was also presented in [23] under the name of statistically grounded aggregation operators. It was shown in [20] that mixture operators can be represented as penalty based functions.

However, as mixture functions are generally not monotone, the conditions on the weighting functions which guarantee monotonicity of the aggregation are important. Several sufficient, but not necessary conditions for monotonicity were established in [19,24]. These conditions are rather restrictive; for example, the weighting function must be monotone. As weak monotonicity has been proposed as a relaxation on the definition of averaging aggregation functions, it makes sense to study sufficient conditions that guarantee the weak monotonicity of mixture functions. A better understanding of this class of functions will support their broader application and this is precisely the goal of this paper. Among other contributions we will identify sufficient conditions for the weak monotonicity of Gini means.

The remainder of this article is structured as follows. In Section 2 we provide the necessary mathematical foundations that underpin aggregation functions and means, which we rely on in subsequent sections. The recent contribution of weak monotonicity and known properties are also presented. In Section 3 we provide new contributions to the understanding of weakly monotone averaging aggregation functions by providing further properties. Within Section 4 we examine several non-monotone mixture functions and prove conditions under which they are weakly monotone. In Section 5 we present sufficient conditions for the weak monotonicity of Gini means. Our conclusions are presented in Section 6.

2. Preliminaries

2.1. Aggregation functions

In this article we make use of the following notations and assumptions. Without loss of generality we assume that the domain of interest is any closed, non-empty interval $\mathbb{I} = [a,b] \subseteq \bar{\mathbb{R}} = [-\infty, \infty]$ and that tuples in $\mathbb{I}^n$ are defined as $\mathbf{x} = (x_{i,n} | n \in \mathbb{N}, i \in \{1,\ldots,n\})$. We write x_i as the shorthand for $x_{i,n}$ such that it is implicit that $i \in \{1,\ldots,n\}$. Furthermore, $\mathbb{I}^n$ is ordered such that for $\mathbf{x},\mathbf{y} \in \mathbb{I}^n, \mathbf{x} \leqslant \mathbf{y}$ implies that each component of $\mathbf{x}$ is no greater than the corresponding component of $\mathbf{y}$. Unless otherwise stated, a constant vector given as $\mathbf{a}$ is taken to mean
$\mathbf{a} = a\underbrace{(1,1,\ldots,1)}_{n \text{ times}} = a\mathbf{1}$, $a \in \mathbb{R}$, where n is implicit within the context of use.

The vector $\mathbf{x}_\nearrow$ denotes the result of permuting the vector $\mathbf{x}$ such that its components are in non-decreasing order, that is, $\mathbf{x}_\nearrow = \mathbf{x}_\sigma$, where σ is the permutation such that $x_{\sigma(1)} \leqslant x_{\sigma(2)} \leqslant \ldots \leqslant x_{\sigma(n)}$. Similarly, the vector $\mathbf{x}_\searrow$ denotes the result of permuting $\mathbf{x}$ such that $x_{\sigma(1)} \geqslant x_{\sigma(2)} \geqslant \cdots \geqslant x_{\sigma(n)}$. We will make use of the common shorthand notation for a sorted vector, being $\mathbf{x}_{()} = (x_{(1)}, x_{(2)}, \ldots, x_{(n)})$. In such cases the ordering will be stated explicitly and then $x_{(k)}$ represents the k-th largest or smallest element of $\mathbf{x}$ accordingly.

Consider now the following definitions, adopted from [9,15,16,29].

Definition 1. A function $F : \mathbb{I}^n \to \bar{\mathbb{R}}$ is **monotone** (increasing) if $\forall \mathbf{x},\mathbf{y} \in \mathbb{I}^n, \mathbf{x} \leqslant \mathbf{y}$ then $F(\mathbf{x}) \leqslant F(\mathbf{y})$.

Note that we use the terms *increasing/decreasing* not in the strict sense. When the inequality $F(\mathbf{x}) < F(\mathbf{y})$ is strict, we will use the terms *strictly* increasing.

For differentiable functions, monotonicity is equivalent to the condition that the directional derivative $D_{\mathbf{e}_i}(F)(\mathbf{x}) = \nabla F(\mathbf{x}) \cdot \mathbf{e}_i \geqslant 0$ at each point $\mathbf{x} \in \mathbb{I}^n$, for all $i \in \{1,2,\ldots,n\}$, where vectors $\mathbf{e}_i$ come from the canonical Euclidean basis.

Definition 2. A function $F : \mathbb{I}^n \to \mathbb{I}$ *is an* **aggregation function** in $\mathbb{I}^n$ if F is monotone increasing in $\mathbb{I}$ and $F(\mathbf{a}) = a, F(\mathbf{b}) = b$.

Thus the two fundamental properties defining an aggregation function are monotonicity with respect to all arguments and bounds preservation. Further properties of aggregation functions relevant within this article are as follows.

Definition 3. A function F is **idempotent** if for every input $\mathbf{x} = (t,t,\ldots,t), t \in \mathbb{I}$ the output is $F(\mathbf{x}) = t$.

The functions of most interest in this article are those that have averaging behavior.

Definition 4. A function F has **averaging behavior** (or is averaging) if for every $\mathbf{x} \in \mathbb{I}^n$ it satisfies $\min(\mathbf{x}) \leqslant F(\mathbf{x}) \leqslant \max(\mathbf{x})$.