axioms

Article

Least Squares in a Data Fusion Scenario via Aggregation Operators

Gildson Queiroz de Jesus * and Eduardo Silva Palmeira

Postgraduate Program in Science and Technology Computational Modeling (PPGMC), Department of Exact Science and Technology (DCET), State University of Santa Cruz (UESC), Ilhéus 45662-900, BA, Brazil
* Correspondence: gqjesus@uesc.br

Abstract: In this paper, appropriate least-squares methods were developed to operate in data fusion scenarios. These methods generate optimal estimates by combining measurements from a finite collection of samples. The aggregation operators of the average type, namely, ordered weighted averaging (OWA), Choquet integral, and mixture operators, were applied to formulate the optimization problem. Numerical examples about fitting curves to a given set of points are provided to show the effectiveness of the proposed algorithms.

Keywords: least squares; aggregation operators; data fusion

MSC: 93E24; 03E72; 47S40; 94A16

check for updates

Citation: de Jesus, G.Q.; Palmeira, E.S. Least Squares in a Data Fusion Scenario via Aggregation Operators. *Axioms* **2022**, *11*, 678. https://doi.org/10.3390/axioms11120678

Academic Editor: Oscar Castillo

Received: 11 October 2022
Accepted: 25 November 2022
Published: 28 November 2022

1. Introduction

Several studies have been carried out on data science. Datasets play an important role in several areas of knowledge, since information can be extracted from them. This information can be used, for example, in decision making, product improvement, process automation, and trend forecasting [1–3].

A number of methods and algorithms have been developed in the literature to extract different information from datasets through mathematical and computational methods. In general, these algorithms were developed to model datasets collected from a single source. In this regard, few algorithms have been formulated to solve the problem in a data fusion scenario, that is, in a scenario where data comes from different sources [4].

The least-squares method (LSM) is a widely used technique for data modeling based on the minimization of a quadratic function [4–9]. LSM was initially conceived for modeling data from a single source. In [4], an LSM was developed considering a data fusion situation (LSM-DF), that is, a method considering data from different sources. LSM-DF was designed for weighted data fusion.

From a mathematical point of view, the LSM-DF is based on a weighted average of the length of residual vectors of the equations $b_k = A_k x + v_k$ with $k = 1, 2, \ldots, L$, expressed by

$$\sum_{k=1}^{L} ||v_k||^2_{W_k} = \sum_{k=1}^{L} v_k^T W_k v_k$$

where W_k are the weights, that is, an aggregation of L values with their corresponding weightings. Here, a very interesting question arises: is weighted averaging the best method for aggregating the data in all scenarios? Within this context, the study of different aggregation methods has recently gained prominence.

Aggregation operators constitute a subarea of fuzzy theory that has the characteristic of combining finite datasets of the same nature into a single dataset [1,2,6,7,10–19]. These operators are basically classified into three categories: mean, conjunctive, and disjunctive.

Applications of these operators can be found in medical problems, image processing, decision making, and engineering problems.

W_k weights are directly related to the $||v_k||^2$ length of residual vectors. However, in some situations, it would be interesting to dynamically allocate the weights to the W_k weightings, putting more weight on the more important $||v_k||^2$ values. Thus, considering the above, the aggregation operators can be considered to bw a viable alternative to change the behavior of LSM-DF.

This study seeks to optimally combine the least-squares method and the aggregation operators of the average type, more specifically, the ordered weighted averaging (OWA) [3,20–22] Choquet integral, [23,24], and mixture [25,26] operators. Furthermore, the aim of this study is to formulate and solve appropriate least-squares methods to model finite collections of datasets of the same nature. An important goal of these algorithms is to generate optimal estimates that aggregate data of different sources. This is necessary for situations that involve systems that can operate under different failure conditions. A numerical example is presented to show the effectiveness of the proposed algorithm.

This paper is organized as follows: in Section 2, preliminary results are related with an admissible order for matrices, aggregation operators, and LSM. In Section 3, LSM-DF via aggregation operators are deduced. In Section 4, a numerical example is shown.

2. Preliminaries

This section addresses topics that form the theoretical basis for the development of LSM-DF via aggregation operators. Initially, the admissible order for matrices is discussed, followed by the aggregation operators of the average type and the classical least-squares method.

2.1. Admissible Order for Matrices

In this section, we present the concept of admissible order for matrices based on [2,16,27]. This is a special way to consider total orders on the set of all matrices of order $m \times n$ with scalar in $\mathcal{R}$ (set of real numbers) denoted by $\mathcal{R}^{m\times n}$.

Let $A, B \in \mathcal{R}^{m\times n}$. It is clear that $A \leq_M B$ given by

$$A \leq_M B \ \text{ if and only if } \ a_{ij} \leq b_{ij}, \forall i, j$$

is a partial order on $\mathcal{R}^{m\times n}$.

Considering a matrix $A \in \mathcal{R}^{m\times n}$ as a vector of columns, i.e., $A = [A_1, A_2, \ldots, A_n]$ where A_i are the columns of A ($i \in \{1, 2, \ldots, n\}$), then $\leq$ can be defined as

$$A \leq_M B \ \text{ if and only if } \ A_i \leq_M B_i, \forall i \in \{1, 2, \ldots, n\}.$$

One can extend that partial order for a total order by considering the concept of admissible order as follows.

Definition 1. *A total order $\preccurlyeq$ on $\mathcal{R}^{m\times n}$ is admissible if, for each $A, B \in \mathcal{R}^{m\times n}$ we have that $A \preccurlyeq B$ whenever $A \leq_M B$.*

Example 1. *Let be A and B column matrices on $\mathcal{R}^{m\times 1}$ and $\pi_i(A) = a_{i1}$ the projection on the i-th line of A. Then,*

$$A \preccurlyeq_c B \Leftrightarrow \exists k \in \{1, 2, \ldots, m\} \text{ s.t. } \pi_k(A) < \pi_k(B) \text{ and } \forall i, 1 \leq i < k, \pi_i(A) = \pi_i(B)$$

is an admissible order.

Therefore, one can generalize an admissible order on $\mathcal{R}^{m\times n}$ by considering the following definition: Let $A, B \in \mathcal{R}^{m\times n}$ such that $A = [A_1, A_2, \ldots, A_n]$ and $B = [B_1, B_2, \ldots, B_n]$. Then

$$A \preccurlyeq_M B \Leftrightarrow \exists k \in \{1, 2, \ldots, m\} \text{ s.t. } A_k \preccurlyeq_c B_k \text{ and } \forall i, 1 \leq i < k, A_i = B_i$$

is an admissible order on $\mathcal{R}^{m\times n}$.

2.2. Aggregation Operators

Aggregation operators are numeric operators that combine multiple input values into a single output value. In this data fusion process, operators aggregate data from different sources to obtain a single unit of data from the conducted analysis. Next, the operators used in this study are presented: OWA, Choquet integral, and mixture operators.

Definition 2 ([12])**.** *(OWA operator) Providing an n- dimensional weight vector, that is, a $W = (w_1, w_2, \ldots, w_n)$ with $\sum_{k=1}^{n} w_k = 1$, the $OWA_W : [0,1]^n \to [0,1]$ function is defined by*

$$OWA_W(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{n} w_k x_{(k)} \tag{1}$$

where $(x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ is the descending order of vector $(x_1, x_2, \ldots, x_n)$ and is named an ordered weighted average function.

Example 2. *Defining the $w = (w_1, w_2, \ldots, w_n)$ vector of weights, where $w_i = 0$ and $w_k = 1$, for some fixed $k \in \{1, 2, \ldots, n\}$. So $OWA_w(x) = x_k$ is the so-called static OWA operator.*

Remark 1. *As one can see in Definition 2, the sum of all the weights in the OWA aggregation results is 1 ($\sum_{k=1}^{n} w_k = 1$). If the weights are matrices, the sum is given by $\sum_{k=1}^{L} ||W_k||_1 = 1$ where $|| \bullet ||_1$ is the norm of the matrices given by*

$$||A||_1 = max_{1\le j\le s} \sum_{i=1}^{r} |a_{ij}|, \text{ where } A \in \mathcal{R}^{rxs}. \tag{2}$$

Remark 2. *The entries in the OWA aggregation must be sorted; if the entries are a matrix, an ordering relation must be used over the set $\mathcal{R}^{mxn}$. So, we can consider an admissible order on $\mathcal{R}^{mxn}$ as defined in 1.*

The next definition is the fuzzy discrete measure, a significant result for the definition of the Choquet integral operator.

Definition 3 ([15])**.** *A discrete fuzzy measure is a function $\mu : 2^{\mathcal{N}} \to [0,1]$ where $\mathcal{N} = \{1, 2, \ldots, n\}$ and $2^{\mathcal{N}}$ is the group of parts of $\mathcal{N}$, such that:*

- *$M_1 : \mu(X) \le \mu(Y)$ when $X \subseteq Y$*
- *$M_2 : \mu(\emptyset) = 0$ and $\mu(\mathcal{N}) = 1$.*

Definition 4 ([10])**.** *(Choquet integral operator) $\mu : 2^{\mathcal{N}} \to [0,1]$ is a discrete fuzzy measure. The discrete Choquet integral related to the measure μ is the function $C_\mu : [0,1]^n \to [0,1]$ defined by:*

$$C_\mu(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{n} x_{[k]} \Big[\mu(\{j \in \mathcal{N} : x_j \ge x_{[k]}\}) - \mu(\{j \in \mathcal{N} : x_j \ge x_{[k+1]}\}) \Big] \tag{3}$$

where $(x_{[1]}, x_{[2]}, \ldots, x_{[n]}) = Sort(x_1, x_2, \ldots, x_n)$ is an ascending ordering of the vector $(x_1, x_2, \ldots, x_n)$ and $x_{[n+1]} = 2$ by convention.

The Choquet integral operator can also be calculated with the following simplified expression:

$$C_\mu(x_1, x_2, \ldots, x_n) = \sum_{k=1}^{n} \Big[x_{[k]} - x_{[k-1]} \Big] \mu(G_k) \tag{4}$$

where $x_{[0]} = 0$ and $G_k = \{ [k], [k+1], \ldots, [n] \}$.

Example 3. *Considering fuzzy discrete measure*

$$\mu_{\perp}(X) = \begin{cases} 1, & se\ X = \mathcal{N} \\ 0, & otherwise. \end{cases}$$

Thus, the following Choquet integral can be defined by:

$$C_{\mu_{\perp}}(x_1, x_2, \ldots, x_n) = [x_{[1]} - x_{[0]}]\mu_{\perp}(G_1) + \cdots + [x_{[n]} - x_{[n-1]}]\mu_{\perp}(G_n).$$

$\mu_{\perp}(G_1) = 1$ *and* $\mu_{\perp}(G_i) = 0$ *for the other values of i; therefore, the result is* $C_{\mu_{\perp}}(x_1, x_2, \ldots, x_n) = x_{[1]} = \min(x_1, x_2, \ldots, x_n)$.

Definition 5 ([15])**.** *(Mixture Operator)* $w_1, w_2, \ldots, w_n : [0,1] \to [0, +\infty)$ *are functions called weight functions. The* $MIX_{w_1,w_2,\ldots,w_n} : [0,1]^n \to [0,1]$ *function is defined by:*

$$MIX_{w_1,w_2,\ldots,w_n}(x_1, x_2, \ldots, x_n) = \frac{\sum_{k=1}^{n} w_k(x_k).x_k}{\sum_{k=1}^{n} w_k(x_k)} \tag{5}$$

is called the mixture function associated with the weight functions $w_1, w_2, \ldots, w_n$.

Example 4. *Defining*

$$w_i(x_i) = \begin{cases} \frac{1}{n}, & se\ x_i = 0 \\ x_i, & otherwise. \end{cases}$$

For simplicity, consider $n = 3$*. In this case, considering that*

$$MIX_{w_1,w_2,w_3}(x_1, x_2, x_3) = \begin{cases} 0, & se\ x_1 = x_2 = x_3 = 0 \\ \frac{{x_1}^2 + {x_2}^2 + {x_3}^2}{x_1 + x_2 + x_3}, & otherwise \end{cases}$$

is the mixture function determined by the w_i *weights defined above.*

2.3. Least-Squares Method

LSM is a widely known and applied mathematical optimization method used to solve several problems, including parameter estimation. This method consists of finding an optimal solution to the problem by minimizing the square of a residual vector.

Considering the equation

$$b = Ax + v \tag{6}$$

where $x \in \mathcal{R}^{n\times1}$ is an unknown vector, $A \in \mathcal{R}^{N\times n}$ is a known parameter matrix, $b \in \mathcal{R}^{N\times1}$ is a known vector, and $v \in \mathcal{R}^{N\times1}$ is a vector named residual.

The least-squares problem is to find a solution $\hat{x}$ that minimizes the length of the residual vector, that is, satisfying the following property:

$$||b - A\hat{x}||^2 \leq ||b - Ax||^2 \tag{7}$$

for all $x \in \mathcal{R}^{n\times1}$. The $|| \bullet ||^2$ denotes the square of Euclidean norm

$$||v||^2 = v^T v. \tag{8}$$

Therefore, the solution to the least-squares problem consists of solving the optimization problem

$$min_x J(x) \tag{9}$$

where the functional cost $J(x)$ is given by

$$\begin{aligned} J(x) &= ||b - Ax||^2 \\ &= (b - Ax)^T(b - Ax). \end{aligned} \tag{10}$$

Theorem 1 ([4]). *(Least-Squares Method) If matrix A has full rank, then there is a single optimal solution $\hat{x}$ for least-squares Problem (9) that is given by*

$$\hat{x} = \left[A^T A\right]^{-1}\left[A^T b\right]. \tag{11}$$

Moreover, the resulting minimal value of the cost function can be written as

$$J(\hat{x}) = b^T b - b^T A\left(A^T A\right)^{-1} A^T b. \tag{12}$$

3. LSM-DF via Aggregation Operators

In this section, LSM-DF is developed via aggregation operators. LSM-DF via an OWA operator, LSM-DF via a Choquet integral operator, and LSM-DF via a mixture operator are also presented. These LSM-DFs are an alternative to estimation problems in the case of several datasources.

The next result is necessary to the proof of the LSM-DF via aggregation operators.

Lemma 1. *If matrices A_k have full rank and matrix W_k is symmetric definite-positive with $k = 1, 2, \ldots, L$, then $\overline{A}^T\overline{WA}$ where*

$$\overline{A} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_L \end{bmatrix}, \overline{W} = \begin{bmatrix} W_1 & 0 & \ldots & 0 \\ 0 & W_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & W_L \end{bmatrix} \tag{13}$$

is nonsingular.

Proof. Let suppose that $\overline{A}^T\overline{WA}$ is singular; then, there must exist a nonzero vector λ, such that $\overline{A}^T\overline{WA}\lambda = 0$, which implies that $\lambda^T\overline{A}^T\overline{WA}\lambda = 0$, i.e.,

$$\lambda^T \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_L \end{bmatrix}^T \begin{bmatrix} W_1 & 0 & \ldots & 0 \\ 0 & W_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & W_L \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_L \end{bmatrix} \lambda = 0 \tag{14}$$

$$\lambda^T A_1^T W_1 A_1 \lambda + \lambda^T A_2^T W_2 A_2 \lambda + \ldots + \lambda^T A_L^T W_L A_L \lambda = 0 \tag{15}$$

(15) can be rewritten as

$$||A_1\lambda||^2_{W_1} + ||A_2\lambda||^2_{W_2} + \ldots + ||A_L\lambda||^2_{W_L} = 0. \tag{16}$$

$|| \bullet ||^2_W$ denotes the square of the weighted Euclidean norm

$$||v||^2_W = v^T W v. \tag{17}$$

As matrices W_k are symmetric definite-positive, it follows from (16) that $||A_k\lambda||^2_{W_k} = 0$ so that $A_k\lambda = 0$ with $k = 1, 2, \ldots, L$. This, in turn, means that the columns of A_k are linearly dependent. Hence, A_k is not full-rank. □

3.1. LSM-DF via OWA Operator

For the deduction of LSM-DF via OWA operator, the following equations should be considered

$$b_{(k)} = A_{(k)}x + v_{(k)}, \quad k = 1, 2, \ldots, L \tag{18}$$

where $x \in \mathcal{R}^{n\times 1}$ is an unknown vector, $A_{(k)} \in \mathcal{R}^{N\times n}$ known parameters arrays, $b_{(k)} \in \mathcal{R}^{N\times 1}$ known vectors, and $v_{(k)} \in \mathcal{R}^{N\times 1}$ vectors named residuals.

A solution to the least-squares problem via operator OWA $\hat{x}$ must minimize the length of the residual vector, that is, it must satisfy the following property:

$$\sum_{k=1}^{L} ||b_{(k)} - A_{(k)}\hat{x}||^2_{W_k} \leq \sum_{k=1}^{L} ||b_{(k)} - A_{(k)}x||^2_{W_k} \tag{19}$$

for all $x \in \mathcal{R}^{n\times 1}$ and where W_k are a positive-definite symmetric matrices.

Optimal solution $\hat{x}$ is found by solving the following minimization problem:

$$min_x \mathcal{J}_{OWA}(x). \tag{20}$$

Functional $\mathcal{J}_{OWA}(x)$ can be defined as

$$\mathcal{J}_{OWA}(x) := OWA_W\Big(J_1(x), J_2(x), \ldots, J_L(x)\Big) \tag{21}$$

where $W = (W_1, W_2, \ldots, W_n)$ are weight matrices and

$$\begin{aligned} J_k(x) &:= ||v_{(k)}||^2 \\ &= ||b_{(k)} - A_{(k)}x||^2, \quad k = 1, 2, \ldots, L. \end{aligned} \tag{22}$$

Therefore, by defining the OWA operator, Function (21) can be rewritten as

$$\begin{aligned} \mathcal{J}_{OWA}(x) &:= \sum_{k=1}^{L} ||b_{(k)} - A_{(k)}x||^2_{W_k} \\ &= \sum_{k=1}^{L} \Big(b_{(k)} - A_{(k)}x\Big)^T W_k \Big(b_{(k)} - A_{(k)}x\Big). \end{aligned} \tag{23}$$

The next theorem brings the solution to the least-squares problem via the OWA operator in (20).

Theorem 2. *(LSM-DF via OWA Operator) If matrices $A_{(k)}$ with $k = 1, 2, \ldots, L$ have full rank and W_k are symmetric definite-positive matrices, then there is a unique optimal solution $\hat{x}$ to the least-squares problem via OWA operator (LSM-DF via OWA operator) that is given by:*

$$\hat{x} = \left[\sum_{k=1}^{L} A^T_{(k)} W_k A_{(k)}\right]^{-1} \left[\sum_{k=1}^{L} A^T_{(k)} W_k b_{(k)}\right]. \tag{24}$$

The corresponding minimal value of $\mathcal{J}_{OWA}(x)$ is

$$\mathcal{J}_{OWA}(\hat{x}) = \sum_{k=1}^{L} b^T_{(k)} W_k b_{(k)} - \sum_{k=1}^{L} b^T_{(k)} W_k A_{(k)} \Big(\sum_{k=1}^{L} A^T_{(k)} W_k A_{(k)}\Big)^{-1} \sum_{k=1}^{L} A^T_{(k)} W_k b_{(k)}. \tag{25}$$

Proof. Consider the cost function

$$\mathcal{J}_{OWA}(x) = \sum_{k=1}^{L} \Big(b_{(k)} - A_{(k)}x\Big)^T W_k \Big(b_{(k)} - A_{(k)}x\Big) \tag{26}$$

$$\begin{aligned}\mathcal{J}_{OWA}(x) \quad &= \quad \left(b_{(1)} - A_{(1)}x\right)^T W_1 \left(b_{(1)} - A_{(1)}x\right) + \left(b_{(2)} - A_{(2)}x\right)^T W_2 \left(b_{(2)} - A_{(2)}x\right) \\ &+ \ldots + \quad \left(b_{(L)} - A_{(L)}x\right)^T W_L \left(b_{(L)} - A_{(L)}x\right)\end{aligned} \tag{27}$$

$$\mathcal{J}_{OWA}(x) = \begin{bmatrix} \left(b_{(1)} - A_{(1)}x\right) \\ \left(b_{(2)} - A_{(2)}x\right) \\ \vdots \\ \left(b_{(L)} - A_{(L)}x\right) \end{bmatrix}^T \begin{bmatrix} W_1 & 0 & \ldots & 0 \\ 0 & W_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & W_L \end{bmatrix} \begin{bmatrix} \left(b_{(1)} - A_{(1)}x\right) \\ \left(b_{(2)} - A_{(2)}x\right) \\ \vdots \\ \left(b_{(L)} - A_{(L)}x\right) \end{bmatrix} \tag{28}$$

$$\begin{aligned}\mathcal{J}_{OWA}(x) \quad &= \quad \left(\begin{bmatrix} b_{(1)} \\ b_{(2)} \\ \vdots \\ b_{(L)} \end{bmatrix} - \begin{bmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(L)} \end{bmatrix} x \right)^T \begin{bmatrix} W_1 & 0 & \ldots & 0 \\ 0 & W_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & W_L \end{bmatrix} \\ &\times \quad \left(\begin{bmatrix} b_{(1)} \\ b_{(2)} \\ \vdots \\ b_{(L)} \end{bmatrix} - \begin{bmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(L)} \end{bmatrix} x \right)\end{aligned} \tag{29}$$

that can be rewritten in matrix form as

$$\mathcal{J}_{OWA}(x) = \left(\overline{b} - \overline{A}x\right)^T \overline{W} \left(\overline{b} - \overline{A}x\right) \tag{30}$$

where

$$\overline{A} = \begin{bmatrix} A_{(1)} \\ A_{(2)} \\ \vdots \\ A_{(L)} \end{bmatrix}, \overline{b} = \begin{bmatrix} b_{(1)} \\ b_{(2)} \\ \vdots \\ b_{(L)} \end{bmatrix}, \overline{W} = \begin{bmatrix} W_1 & 0 & \ldots & 0 \\ 0 & W_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & W_L \end{bmatrix}. \tag{31}$$

Entries $(A_{(1)}, A_{(2)}, \ldots, A_{(L)})$ and $(b_{(1)}, b_{(2)}, \ldots, b_{(L)})$ are descending orders of $(A_1, A_2, \ldots, A_L)$ and $(b_1, b_2, \ldots, b_L)$, respectively. $\overline{W}$ is a diagonal positive-definite symmetric matrix with entries W_k.

To find the critical point in x, $\mathcal{J}_{OWA}(x)$ must be differentiated and equal to zero

$$\begin{aligned} &\frac{\partial}{\partial x}\left[x^T\overline{A}^T\overline{WA}x - x^T\overline{A}^T\overline{Wb} - \overline{b}^T\overline{WA}x + \overline{b}^T\overline{Wb}\right] = 0 \\ \Rightarrow \quad &x^T\overline{A}^T\overline{WA} - \overline{b}^T\overline{WA} = 0. \end{aligned} \tag{32}$$

Via Lemma 1, matrix $\overline{A}^T\overline{WA}$ is invertible. Therefore,

$$\hat{x} = \left[\overline{A}^T\overline{WA}\right]^{-1}\left[\overline{A}^T\overline{Wb}\right]. \tag{33}$$

Replacing (31) into (33), the solution can be rewritten as

$$\hat{x} = \left[\sum_{k=1}^{L} A_{(k)}^T W_k A_{(k)}\right]^{-1}\left[\sum_{k=1}^{L} A_{(k)}^T W_k b_{(k)}\right]. \tag{34}$$

In fact, for the Hermitian matrix to be defined as positive

$$\frac{\partial^2 \mathcal{J}_{OWA}(x)}{\partial x^T \partial x} = \overline{A}^T \overline{WA} = \sum_{k=1}^{L} A_{(k)}^T W_k A_{(k)} > 0 \tag{35}$$

$\mathcal{J}_{OWA}(x)$ in (30) must be a strictly convex function; therefore, $\hat{x}$ is a unique global minimum.

The minimal cost $\mathcal{J}_{OWA}(\hat{x})$ can be expressed as

$$\begin{aligned} \mathcal{J}_{OWA}(\hat{x}) &= \sum_{k=1}^{L} ||b_{(k)} - A_{(k)}\hat{x}||_{W_k}^2 \\ &= \left(\overline{b} - \overline{A}\hat{x}\right)^T \overline{W}\left(\overline{b} - \overline{A}\hat{x}\right) \\ &= \overline{b}^T \overline{Wb} - \overline{b}^T \overline{WA}\hat{x} - \hat{x}^T \overline{A}^T \overline{Wb} + \hat{x}^T \overline{A}^T \overline{WA}\hat{x} \end{aligned} \tag{36}$$

Replacing (33) into (36) results in

$$\mathcal{J}_{OWA}(\hat{x}) = \overline{b}^T \overline{Wb} - \overline{b}^T \overline{WA}\left(\overline{A}^T \overline{WA}\right)^{-1} \overline{A}^T \overline{Wb}. \tag{37}$$

Replacing (31) into (37), the optimal cost can be rewritten as

$$\mathcal{J}_{OWA}(\hat{x}) = \sum_{k=1}^{L} b_{(k)}^T W_k b_{(k)} - \sum_{k=1}^{L} b_{(k)}^T W_k A_{(k)} \left(\sum_{k=1}^{L} A_{(k)}^T W_k A_{(k)}\right)^{-1} \sum_{k=1}^{L} A_{(k)}^T W_k b_{(k)}. \tag{38}$$

□

Remark 3. *Applying $k = 1$ in Theorem (2), the LSM-DF via OWA operator reduces to the classical LSM in Theorem (1).*

3.2. LSM-DF via Choquet Integral Operator

The deduction of the LSM-DF via the Choquet integral operator follows from the equations

$$b_{[k]} = A_{[k]}x + v_{[k]}, \quad k = 1, 2, \ldots, L \tag{39}$$

where $x \in \mathcal{R}^{n\times 1}$ is an unknown vector, $A_{[k]} \in \mathcal{R}^{N\times n}$ known parameters matrices, $b_{[k]} \in \mathcal{R}^{N\times 1}$ known vectors, and $v_{[k]} \in \mathcal{R}^{N\times 1}$ vectors named residuals.

A solution to the least-squares problem via the Choquet integral operator $\hat{x}$ must minimize the length of the residual vector, that is, it must satisfy the following property:

$$\sum_{k=1}^{L} ||b_{[k]} - A_{[k]}\hat{x}||_{I\mu(G_k)}^2 \leq \sum_{k=1}^{L} ||b_{[k]} - A_{[k]}x||_{I\mu(G_k)}^2 \tag{40}$$

for all $x \in \mathcal{R}^{n\times 1}$ and where $I\mu(G_k)$ is a matrix identity multiplied by discrete fuzzy measure.

The optimal solution $\hat{x}$ is found by solving the following minimization problem:

$$min_x \mathcal{J}_{C_\mu}(x) \tag{41}$$

Functional $\mathcal{J}_{C_\mu}(x)$ can be defined as

$$\mathcal{J}_{C_\mu}(x) := C_\mu\left(\bar{J}_1(x), \bar{J}_2(x), \ldots, \bar{J}_L(x)\right) \tag{42}$$

where

$$\begin{aligned} \bar{J}_k(x) &:= ||v_{[k]} - v_{[k-1]}||^2 \\ &= ||(b_{[k]} - A_{[k]}x) - (b_{[k-1]} - A_{[k-1]}x)||^2 \\ &= ||\Big(b_{[k]} - b_{[k-1]}\Big) - \Big(A_{[k]} - A_{[k-1]}\Big)x||^2, \quad k = 1, 2, \ldots, L. \end{aligned} \tag{43}$$

Therefore, by defining the Choquet integral operator, Function (42) can be rewritten as

$$\begin{aligned} \mathcal{J}_{C_\mu}(x) &:= \sum_{k=1}^{L} ||\Big(b_{[k]} - b_{[k-1]}\Big) - \Big(A_{[k]} - A_{[k-1]}\Big)x||^2_{I\mu(G_k)} \\ &= \sum_{k=1}^{L} \Big[\Big(b_{[k]} - b_{[k-1]}\Big) - \Big(A_{[k]} - A_{[k-1]}\Big)x\Big]^T I\mu(G_k) \\ &\bullet \Big[\Big(b_{[k]} - b_{[k-1]}\Big) - \Big(A_{[k]} - A_{[k-1]}\Big)x\Big]. \end{aligned} \tag{44}$$

where $I\mu(G_k)$ is a positive-definite symmetric matrix.

The next theorem brings the solution to the least-squares problem via the Choquet integral operator in (41).

Theorem 3. *(LSM-DF via Choquet Integral Operator) If the $A_{[k]} - A_{[k-1]}$ matrices with $k = 1, 2, \ldots, L$ have a full rank and $I\mu(G_k)$ are symmetric definite-positive matrices, then there is a single optimal solution $\hat{x}$ for the least-squares problem via Choquet integral operator (LSM-DF via Choquet integral operator) that is given by:*

$$\begin{aligned} \hat{x} &= \left[\sum_{k=1}^{L} \Big(A_{[k]} - A_{[k-1]}\Big)^T I\mu(G_k) \Big(A_{[k]} - A_{[k-1]}\Big)\right]^{-1} \\ &\bullet \left[\sum_{k=1}^{L} \Big(A_{[k]} - A_{[k-1]}\Big)^T I\mu(G_k) \Big(b_{[k]} - b_{[k-1]}\Big)\right]. \end{aligned} \tag{45}$$

The corresponding minimal value of $\mathcal{J}_{C_\mu}(x)$ is

$$\begin{aligned} \mathcal{J}_{C_\mu}(\hat{x}) &= \sum_{k=1}^{L} \Big(b_{[k]} - b_{[k-1]}\Big)^T I\mu(G_k) \Big(b_{[k]} - b_{[k-1]}\Big) \\ &- \sum_{k=1}^{L} \Big(b_{[k]} - b_{[k-1]}\Big)^T I\mu(G_k) \Big(A_{[k]} - A_{[k-1]}\Big) \\ &\bullet \left(\sum_{k=1}^{L} \Big(A_{[k]} - A_{[k-1]}\Big)^T I\mu(G_k) \Big(A_{[k]} - A_{[k-1]}\Big)\right)^{-1} \\ &\bullet \sum_{k=1}^{L} \Big(A_{[k]} - A_{[k-1]}\Big)^T I\mu(G_k) \Big(b_{[k]} - b_{[k-1]}\Big). \end{aligned} \tag{46}$$

Proof. Consider functional cost

$$\begin{aligned} \mathcal{J}_{C_\mu}(x) &= \sum_{k=1}^{L} \Big[\Big(b_{[k]} - b_{[k-1]}\Big) - \Big(A_{[k]} - A_{[k-1]}\Big)x\Big]^T I\mu(G_k) \\ &\bullet \Big[\Big(b_{[k]} - b_{[k-1]}\Big) - \Big(A_{[k]} - A_{[k-1]}\Big)x\Big] \end{aligned} \tag{47}$$

Using the matrices, this can be rewritten as

$$\mathcal{J}_{C_\mu}(x) = (b' - A'x)^T W' (b' - A'x) \tag{48}$$

where

$$A' = \begin{bmatrix} A_{[1]} - A_{[0]} \\ A_{[2]} - A_{[1]} \\ \vdots \\ A_{[L]} - A_{[L-1]} \end{bmatrix}, b' = \begin{bmatrix} b_{[1]} - b_{[0]} \\ b_{[2]} - b_{[1]} \\ \vdots \\ b_{[L]} - b_{[L-1]} \end{bmatrix},$$

$$W' = \begin{bmatrix} I\mu(G_1) & 0 & \dots & 0 \\ 0 & I\mu(G_2) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & I\mu(G_L) \end{bmatrix}. \tag{49}$$

Entries $(A_{[1]}, A_{[2]}, \dots, A_{[L]})$, and $(b_{[1]}, b_{[2]}, \dots, b_{[L]})$ are ascending orders of $(A_1, A_2, \dots, A_L)$, and $(b_1, b_2, \dots, b_L)$, respectively. W' is a diagonal symmetric definite-positive matrix with entries $I\mu(G_k)$.

On the basis of Function (48) and the solution of LSM-DF via the OWA operator presented in Theorem (2), the solution to Optimization Problem (41) is given by

$$\hat{x} = \left[A'^T W' A'\right]^{-1} \left[A'^T W' b'\right]. \tag{50}$$

which, through Matrices (49), can be rewritten as

$$\hat{x} = \left[\sum_{k=1}^{L} \left(A_{[k]} - A_{[k-1]}\right)^T I\mu(G_k) \left(A_{[k]} - A_{[k-1]}\right)\right]^{-1} \bullet \left[\sum_{k=1}^{L} \left(A_{[k]} - A_{[k-1]}\right)^T I\mu(G_k) \left(b_{[k]} - b_{[k-1]}\right)\right]. \tag{51}$$

Similar to the procedure performed in Theorem (2), the minimal cost $\mathcal{J}_{C_\mu}(\hat{x})$ can be expressed as

$$\mathcal{J}_{C_\mu}(\hat{x}) = b'^T W' b' - b'^T W' A' \left(A'^T W' A'\right)^{-1} A'^T W' b'. \tag{52}$$

Replacing (49) into (52), the optimal cost can be rewritten as

$$\begin{aligned} \mathcal{J}_{C_\mu}(\hat{x}) = & \sum_{k=1}^{L} \left(b_{[k]} - b_{[k-1]}\right)^T I\mu(G_k) \left(b_{[k]} - b_{[k-1]}\right) \\ - & \sum_{k=1}^{L} \left(b_{[k]} - b_{[k-1]}\right)^T I\mu(G_k) \left(A_{[k]} - A_{[k-1]}\right) \\ \bullet & \left(\sum_{k=1}^{L} \left(A_{[k]} - A_{[k-1]}\right)^T I\mu(G_k) \left(A_{[k]} - A_{[k-1]}\right)\right)^{-1} \\ \bullet & \sum_{k=1}^{L} \left(A_{[k]} - A_{[k-1]}\right)^T I\mu(G_k) \left(b_{[k]} - b_{[k-1]}\right). \end{aligned} \tag{53}$$

□

Remark 4. *$A_{[0]}$ is the null matrix and $b_{[0]}$ is the null vector by convention.*

Remark 5. *By applying $k = 1$ in Theorem (3), the LSM-DF via Choquet integral operator reduces to the classical LSM in Theorem (1).*

3.3. LSM-DF via Mixture Operator

For the deduction of the LSM-DF via the mixture operator, it is necessary to adapt the mixture operator presented in Definition (5).

The weight functions that are dynamic in the mixture operator uses were previously calculated and became constant (static) weight functions. Thus, the adapted mixture operator is calculated in two steps. In the first step, the weights are calculated and fixed. In the next step, aggregations are carried out. The next definition brings the adapted mixture operator.

Definition 6. *(Adapted Mixture Operator) The adapted MIX function can be calculated using the following steps:*

- *Step 1: weight functions $w_k(x_k)$ with $k = 1,,2,\ldots,n$ can be calculated and fixed as follows:*

$$w_1(x_1) = w_1, w_2(x_2) = w_2, \ldots, w_n(x_n) = w_n. \tag{54}$$

- *Step 2: with the fixed weight functions, the MIX function can be calculated as follows:*

$$MIX_{w_1,w_2,\ldots,w_n}(x_1, x_2, \ldots, x_n) = \frac{\sum_{k=1}^{n} w_k x_k}{\sum_{k=1}^{n} w_k}. \tag{55}$$

Now, the LSM-DF via the mixture operator must be deduced. The following equation must be considered:

$$b_k = A_k x + v_k, \quad k = 1, 2, \ldots, L \tag{56}$$

where $x \in \mathcal{R}^{n\times 1}$ is an unknown vector, $A_k \in \mathcal{R}^{N\times n}$ known parameters matrices, $b_k \in \mathcal{R}^{N\times 1}$ known vectors, and $v_k \in \mathcal{R}^{N\times 1}$ vectors named residuals.

A solution to the least-squares problem via the mixture operator must minimize the length of the residual vector, that is, it must satisfy the following property:

$$\frac{\sum_{k=1}^{L} ||b_k - A_k\hat{x}||^2_{W_k}}{\sum_{k=1}^{L} ||W_k||^2} \leq \frac{\sum_{k=1}^{L} ||b_k - A_k x||^2_{W_k}}{\sum_{k=1}^{L} ||W_k||^2} \tag{57}$$

for all $x \in \mathcal{R}^{n\times 1}$ and where W_k is a positive-definite symmetric matrix.

Optimal solution $\hat{x}$ is found by solving the following minimization problem:

$$min_x J_{MIX}(x) \tag{58}$$

Functional $\mathcal{J}_{MIX}(x)$ can be defined as

$$\mathcal{J}_{MIX}(x) := MIX_{W_1,W_2,\ldots,W_L}\left(\underline{J}_1(x), \underline{J}_2(x), \ldots, \underline{J}_L(x)\right) \tag{59}$$

where

$$\begin{aligned} \underline{J}_k(x) &:= ||v_k||^2 \\ &= ||b_k - A_k x||^2, \quad k = 1, 2, \ldots, L. \end{aligned} \tag{60}$$

By defining Mixture Operator (59), the function can be rewritten as

$$\begin{aligned} \mathcal{J}_{MIX}(x) &:= \frac{\sum_{k=1}^{L} ||b_k - A_k x||^2_{W_k}}{\sum_{k=1}^{L} ||W_k||^2} \\ &= \frac{\sum_{k=1}^{L} (b_k - A_k x)^T W_k (b_k - A_k x)}{\sum_{k=1}^{L} ||W_k||^2}. \end{aligned} \tag{61}$$

The next theorem brings the solution to the least-squares problem via the mixture operator in (58).

Theorem 4. *(LSM-DF via Mixture Operator) If the A_k matrices with $k = 1, 2, \ldots, L$ have a full rank and W_k are symmetric definite-positive matrices, then there is a single optimal solution $\hat{x}$ to the least-squares problem via the mixture operator (LSM-DF via mixture operator)(58) that is given by:*

$$\hat{x} = \left[\sum_{k=1}^{L} A_k^T W_k A_k\right]^{-1} \left[\sum_{k=1}^{L} A_k^T W_k b_k\right]. \tag{62}$$

The corresponding minimal value of $\mathcal{J}_{MIX}(x)$ is

$$\mathcal{J}_{MIX}(\hat{x}) \quad = \quad \sum_{k=1}^{L} b_k^T W_k b_k - \sum_{k=1}^{L} b_k^T W_k A_k \Big(\sum_{k=1}^{L} A_k^T W_k A_k\Big)^{-1} \sum_{k=1}^{L} A_k^T W_k b_k. \tag{63}$$

Proof. Consider the function

$$\mathcal{J}_{MIX}(x) = \left[\frac{\sum_{k=1}^{L}(b_k - A_k x)^T W_k (b_k - A_k x)}{\sum_{k=1}^{L} ||W_k||^2}\right] \tag{64}$$

that can be rewritten as

$$\mathcal{J}_{MIX}(x) = \alpha(\beta - \mathcal{A}x)^T \mathcal{W}(\beta - \mathcal{A}x) \tag{65}$$

where

$$\begin{aligned} \mathcal{A} &= \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_L \end{bmatrix}, \beta = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_L \end{bmatrix}, \mathcal{W} = \begin{bmatrix} W_1 & 0 & \ldots & 0 \\ 0 & W_2 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & W_L \end{bmatrix}, \\ \alpha &= \frac{1}{\sum_{k=1}^{L} ||W_k||^2} \end{aligned} \tag{66}$$

where $\mathcal{W}$ is a diagonal positive-definite symmetric matrix with entries W_k.

To find the solution to optimization problem $\hat{x}$, $J(x)$ must be differentiated in (65) and equal to zero. On the basis of the theorem, the solution of the derivative is given by

$$\begin{aligned} \frac{\partial}{\partial x}\Big[\alpha(\beta - \mathcal{A}x)^T \mathcal{W}(\beta - \mathcal{A}x)\Big] &= 0 \\ \alpha\frac{\partial}{\partial x}\Big[(\beta - \mathcal{A}x)^T \mathcal{W}(\beta - \mathcal{A}x)\Big] &= 0. \end{aligned} \tag{67}$$

On the basis of Theorem (2), the solution of the derivative is given by

$$\alpha\Big[x^T \mathcal{A}^T \mathcal{W}\mathcal{A} - \beta^T \mathcal{W}\mathcal{A}\Big] = 0. \tag{68}$$

Therefore,

$$\hat{x} = \Big[\mathcal{A}^T \mathcal{W}\mathcal{A}\Big]^{-1}\Big[\mathcal{A}^T \mathcal{W}\beta\Big]. \tag{69}$$

Through Matrices (66), the solution can be rewritten as

$$\hat{x} = \left[\sum_{k=1}^{L} A_k^T W_k A_k\right]^{-1} \left[\sum_{k=1}^{L} A_k^T W_k b_k\right]. \tag{70}$$

Minimal cost $\mathcal{J}_{MIX}(\hat{x})$ can be expressed as

$$\mathcal{J}_{MIX}(\hat{x}) = \beta^T \mathcal{W}\beta - \beta^T \mathcal{W}\mathcal{A}\hat{x} - \hat{x}^T \mathcal{A}^T \mathcal{W}\beta + \hat{x}^T \mathcal{A}^T \mathcal{W}\mathcal{A}\hat{x} \tag{71}$$

replacing(69) into (71), the result is

$$\mathcal{J}_{MIX}(\hat{x}) = \beta^T \mathcal{W}\beta - \beta^T \mathcal{W}\mathcal{A}\left(\mathcal{A}^T \mathcal{W}\mathcal{A}\right)^{-1} \mathcal{A}^T \mathcal{W}\beta. \tag{72}$$

Replacing (66) into (72), the optimal cost can be rewritten as

$$\mathcal{J}_{MIX}(\hat{x}) = \sum_{k=1}^{L} b_k^T W_k b_k - \sum_{k=1}^{L} b_k^T W_k A_k \left(\sum_{k=1}^{L} A_k^T W_k A_k\right)^{-1} \sum_{k=1}^{L} A_k^T W_k b_k. \tag{73}$$

□

Remark 6. *The optimal solution of the LSM-DF via a mixture operator reduces to the LSM-DF in [4].*

4. Illustrative Example

In this section, we present artificially created (by authors) datasets in order to illustrate the behavior, effectiveness, and the relationship between the proposed methods for finding the best fitting curve to a given set of points from a mathematical point of view. Table 1 shows two simulated datasets about income and consumption.

Table 1. Simulated datasets about income and consumption.

Income (x_1)	Consumption (y_1)	Income (x_2)	Consumption (y_2)
139	122	140	123
126	114	129	117
90	86	92	89
144	134	145	136
163	146	163	147
136	107	138	109
61	68	64	68
62	117	63	119
41	71	43	73
120	98	122	100

First, the LSM was separately applied to the datasets, and the following results were found:

$$\hat{y}_1 = 0.49x_1 + 52.69, \tag{74}$$

$$\hat{y}_2 = 0.49x_2 + 53.65. \tag{75}$$

The MSEs between $\hat{y}_1$ with y_1 and $\hat{y}_2$ with y_2 were 211.52 and 221.67, respectively. Model (74) was more accurate than Model (75).

Second, the LSM-DF via OWA, Choquet integral, and mixture operators were calculated in the two datasets, and the following weighting matrices were used in the simulation: $W_1 = 0.7 * diag(10)$ and $W_2 = 0.3 * diag(10)$; more weight was given to W_1 than to W_2. The following results were found:

$$\hat{y}_O = 0.49x + 53.34, \tag{76}$$

$$\hat{y}_C = 0.49x + 52.65, \tag{77}$$

$$\hat{y}_M = 0.49x + 52.95. \tag{78}$$

The MSEs between $\hat{y}_O$, $\hat{y}_C$ and $\hat{y}_M$ with y_1 were 211.18, 211.57, 211.28, respectively. The MSEs between $\hat{y}_O$, $\hat{y}_C$ and $\hat{y}_M$ with y_2 were 222.14, 223.87 and 223 respectively. Tables 2 and 3 compare samples with regard to x_1 and x_2, respectively, of Equations (76)–(78). Table 4 compares the samples of y_1 to the samples generated by Equations (74), (76)–(78). Table 5 compares the samples of y_2 with the samples generated with Equations (76)–(78).

MSE shows that Models (76)–(78) were more accurate than Model (74). The LSM-DF via OWA, Choquet integral, and mixture operators outperformed the LSM.

Table 2. Sample with regard to x_1 of Equations (76)–(78).

Income (x_1)	Consumption ($\hat{y}_O$)	Consumption ($\hat{y}_C$)	Consumption ($\hat{y}_M$)
139	121.45	120.76	121.06
126	115.08	114.39	114.69
90	97.44	96.75	97.05
144	123.90	123.21	123.51
163	133.21	132.52	132.82
136	1119.98	119.29	119.59
61	83,23	82.54	82.84
62	83.72	83.03	83.33
41	73.43	72.74	73.04
120	112.14	111.45	111.75

Table 3. Sample with regard to x_2 of Equations (76)–(78).

Income (x_2)	Consumption ($\hat{y}_O$)	Consumption ($\hat{y}_C$)	Consumption ($\hat{y}_M$)
140	121.94	121.25	121.55
129	116.55	115.86	116.16
92	98.42	97.73	98.03
145	124.39	123.70	124
163	133.21	132.52	132.82
138	120.96	120.27	120.57
64	87.40	84.01	84.31
63	84.21	83.52	83.82
43	74.41	73.72	74.02
122	113.12	112.43	112.73

Table 4. Sample of y_1 and samples generated with Equations (74), (76)–(78).

y_1	$\hat{y}_1$	$\hat{y}_O$	$\hat{y}_C$	$\hat{y}_M$
122	120.80	121.45	120.76	121.06
114	114.43	115.08	114.39	114.69
86	96.79	97.44	96.75	97.05
134	123.25	123.90	123.21	123.51
146	132.56	133.21	132.52	132.82
107	119.33	119.98	119.29	119.59
68	82.58	83,23	82.54	82.84
117	83.07	83.72	83.03	83.33
71	72.78	73.43	72.74	73.04
98	111.49	112.14	111.45	111.75

Table 5. Sample of y_2 and the samples generated by Equations (76)–(78).

y_2	$\hat{y}_2$	$\hat{y}_O$	$\hat{y}_C$	$\hat{y}_M$
123	122.25	121.94	121.25	121.55
117	116.86	116.55	115.86	116.16
89	98.73	98.42	97.73	98.03
136	124.70	124.39	123.70	124
147	133.52	133.21	132.52	132.82
109	121.27	120.96	120.27	120.57
68	85.01	87.40	84.01	84.31
119	84.52	84.21	83.52	83.82
73	74.72	74.41	73.72	74.02
100	113.43	113.12	112.43	112.73

5. Conclusions

In this paper, the LSM-DF was studied through aggregation operators in order to explore different ways to aggregate data. More specifically, the LSM-DF via an OWA operator, the LSM-DF via a Choquet integral operator, and the LSM-DF via a mixture operator were defined. These operators were particularly chosen due to their efficiency when applied to other methods in different areas of knowledge [12,13,22,24,26]. These new methods provide a theoretical framework with variations of the classic least square, which may be more suitable in certain applications. For instance, LSM-DF via OWA operator could be chosen for situations where one wants to place greater weights on the first data entries.

The main objective of developing these methods is to estimate an optimal parameter for situations involving more than one dataset, and to show how it can be changed for different types of data. The methods were mathematically demonstrated by applying aggregation operators of the average type to optimization problem. The illustrate example was set up to demonstrate the mathematical behavior of these procedures trough fitting curves in comparison with an approach that does not incorporate the aggregation operators in its formulation.

In future studies, we want to explore some applications that can show the advantages and disadvantages of each method, and set up LSM for other aggregation operators such as a weighted OWA (WOWA) operator and a Sugeno integral operator. Furthermore, these methods will be extended to models subject to parametric uncertainties.

Author Contributions: Conceptualization, G.Q.d.J. and E.S.P.; Methodology, G.Q.d.J. and E.S.P.; Formal analysis, G.Q.d.J. and E.S.P.; Investigation, G.Q.d.J. and E.S.P.; Writing—original draft, G.Q.d.J.; Writing—review & editing, E.S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cheng, C.H.; Wang, J.W.; Wu, M.C. OWA-weighted based clustering method for classification problem. *Expert Syst. Appl.* **2009**, *36*, 4988–4995. [CrossRef]
2. Milfont, T.; Mezzomo, I.; Bedregal, B.; Mansilla, E.; Bustince, H. Aggregation functions on n-dimensional ordered vectors equipped with an admissible order and an application in multi-criteria group decision-making. *Int. J. Approx. Reason.* **2021**, *137*, 34–50. [CrossRef]
3. Flores-Sosa, M.; León-Castro, E.; Merigó, J.M.; Yager, R.R. Forecasting the exchange rate with multiple linear regression and heavy ordered weighted average operators. *Eur. J. Oper. Res.* **2022**, *248*, 108863.
4. Sayed, A.H.; Al-Naffouri, T.Y.; Kailath, T. Robust Estimation for Uncertain Models in a Data Fusion Scenario. *IFAC Proc. Vol.* **2000**, *33*, 899–904. [CrossRef]
5. Kailath, T.; Sayed A.S.; Hassibi, B. *Linear Estimation*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2000; 854p.
6. Sayed, A.H.; Chandrasekaran, S. Parameter estimation with multiple sources and levels of uncertainties. *IEEE Trans. Signal Process.* **2000**, *48*, 680–692. [CrossRef]
7. Lopes, C.G.; Sayed, A.H. Diffusion least-mean squares over adaptive networks: Formulation and performance analysis. *IEEE Trans. Signal Process.* **2008**, *56*, 3122–3136. [CrossRef]
8. Cattivelli, F.; Lopes, C.G.; Sayed, A.H. Diffusion recursive least-squares for distributed estimation over adaptive networks. *IEEE Trans. Signal Process.* **2008**, *56*, 1865–1877. [CrossRef]
9. Takahashi, N.; Yamada, I.; Sayed, A.H. Diffusion least-mean squares with adaptive combiners: Formulation and performance analysis. *IEEE Trans. Signal Process.* **2010**, *58*, 4795–4810. [CrossRef]
10. Choquet, G. Theory of capacities. *Ann. de lÍnstitut Fourier* **1953**, *5*, 131–295. [CrossRef]
11. Give'on, Y. Lattice matrices. *Inf. Control* **1964**, *7*, 477–484. [CrossRef]
12. Yager, R.R. Ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Trans. Syst. Man Cybern.* **1988**, *18*, 183–190. [CrossRef]

13. Zhou, S.M.; Chiclana, F.; John, R.I.; Garibald, J.M. Type-1 OWA operators for aggregating uncertain information with uncertain weights induced by type-2 linguistic quantifiers. *Fuzzy Sets Syst.* **2008**, *159*, 3281–3296. [CrossRef]
14. Paternain, D.; Fernandez, J.; Bustince, H.; Mesiar, R.; Beliakov, G. Construction of image reduction operators using averaging aggregation functions. *Fuzzy Sets Syst.* **2015**, *261*, 87–111. [CrossRef]
15. Beliakov, G.; Bustince, H.; Calvo, T. *A Practical Guide to Averaging Functions (Studies in Fuzziness and Soft Computing)*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 329.
16. Bedregal, B.; Bustince, H.; Palmeira, E.; Dimuro, G.; Fernandez, J. Generalized Interval-valued OWA operators with interval weights derived from interval-valued overlap functions. *Int. J. Approx. Reason.* **2017**, *90*, 1–16. [CrossRef]
17. Joy, G. The Determinant and Rank of a Lattice Matrix. *Glob. J. Pure Appl. Math.* **2017**, *13*, 1745–1761.
18. Dimuro, G.P.; Fernandez, J.; Bedregal, B.; Mesiar, R.; Sanz, J.A.; Lucca, G.; Bustince, H. The state-of-art of the generalization of the Choquet integral: From aggregation and pre-aggregation to ordered directionally monotone functions. *Inf. Fusion* **2020**, *57*, 27–43. [CrossRef]
19. Asmus, T.; Dimuro, G.; Bedregal, B.; Sanz, J.A.; Fernandez, J.; Rodriguez-Martinez, J.; Mesiar, R.; Bustince, H. A constructive framework to define fusion functions with floating domains in arbitrary closed real intervals. *Inf. Sci.* **2022**, *601*, 800–829. [CrossRef]
20. Flores-Sosa, M.; Avilés-Ochoa, E.; Merigó, J.M.; Yager, R.R. Volatility GARCH models with the ordered weighted average (OWA) operators. *Inf. Sci.* **2021**, *565*, 46–61. [CrossRef]
21. Medina, J.; Yager, R.R. OWA operators with functional weights. *Fuzzy Sets Syst.* **2021**, *414*, 38–56. [CrossRef]
22. Flores-Sosa, M.; Avilés-Ochoa, E.; Merigó, J.M.; Kacprzyk, J. The OWA operator in multiple linear regression. *Appl. Soft Comput.* **2022**, *124*, 108985. [CrossRef]
23. Llamazares, B. Constructing Choquet integral-based operators that generalize weighted means and OWA operators. *Inf. Fusion* **2022**, *23*, 131–138. [CrossRef]
24. Jia, X.; Wang, Y. Choquet integral-based intuitionistic fuzzy arithmetic aggregation operators in multi-criteria decision-making. *Expert Syst. Appl.* **2022**, *191*, 116242. [CrossRef]
25. Pereira, R.A.M.; Ribeiro, R.A. Aggregation with generalized mixture operators using weighting functions. *Fuzzy Sets Syst.* **2003**, *137*, 43–58. [CrossRef]
26. Ribeiro, R.A.; Pereira, R.A.M. Generalized mixture operators using weighting functions: A comparative study with WA and OWA. *Eur. J. Oper. Res.* **2003**, *145*, 329–342. [CrossRef]
27. Santana, F.; Bedregal, B.; Viana, P.; Bustince, H. On admissible orders over closed subintervals of $[0,1]$. *Fuzzy Sets Syst.* **2021**, *399*, 44–54. [CrossRef]