



The Component Weighted Median Absolute Deviations Problem

Vedran Novoselac

*Department of Mathematics, Mechanical Engineering Faculty in Slavonski Brod,
University of Slavonski Brod, Trg Ivane Brlić Mažuranić 2, 35000 Slavonski Brod, Croatia*

Abstract. This paper considers the problem of robust modeling by using the well-known Least Absolute Deviation (LAD) regression. For that purpose, the approximation function is designed and analyzed, which is based on a certain component weight of the Weighted Median of Data. It is shown that the proposed approximation function is a piecewise constant function with finitely many pieces with respect to the model parameter. Thereby, an investigation of regions of constant values of the approximation function is conducted. It is established that the designed model based on the Component Weighted Median Absolute Deviations estimates an optimal model parameter on a finite set, which describes the corresponding regions. Furthermore, the specified restriction of the approximation function is observed and analyzed, in order to examine the observed problem.

2020 Mathematics Subject Classifications: 26E60, 26A15, 62J99

Key Words and Phrases: Weighted median, approximation function, robust regression

1. Introduction

This paper deals with robust regression modeling, and thus the well known Least Absolute Deviations (LAD) regression model is considered. In general the LAD regression involves finding estimates which minimize a sum of the residuals' absolute values, which has important applications in many fields, including statistics and numerical analysis [1, 7, 9].

In that sense, the approximation function $F: \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ is modeled, which considers a certain weight of the Weighted Median of Data as a variable. It is well known that the Weighted Median of Data is a robust estimator, which has a great number of applications in many fields of applied research like statistics, data analysis, outlier detection, image processing, etc. [2, 3, 8]. Considering the approximating of a model, the L_1 norm error model function is defined

$$\Delta(w) = \sum_{j=1}^n |y_j - F(\mathbf{x}_j; w)|,$$

DOI: <https://doi.org/10.29020/nybg.ejpam.v13i4.3839>

Email address: Vedran.Novoselac@sfsb.hr (V. Novoselac)

where $Y = \{y_j \in \mathbb{R}: j \in \{1, \dots, n\}\}$ presents the observed target data (dependent variables), $X = \{\mathbf{x}_j \in \mathbb{R}^m: j \in \{1, \dots, n\}\}$ the feature data (independent variables), and $w > 0$ the model parameter [6]. The L_1 norm based models are often not trivial to solve, and thereby an investigation of the Component Weighted Median Function (CWMF) $F: \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ is conducted, where it is shown that the restricted approximation function $F|_{\mathcal{D}}$, $\mathcal{D} = \mathbf{z} \times \mathbb{R}_+$, $\mathbf{z} \in \mathbb{R}^m$, is a piecewise constant function with finitely many pieces with respect to the parameter $w > 0$. According to that, the minimization of the Component Weighted Median Absolute Deviation (CWMAD) model Δ is analyzed, where it is shown that the minimization of Δ , i.e.

$$\min_{w>0} \Delta(w)$$

can be derived on the finite set, which describes the regions of constant values of the restricted CWMF. Furthermore, the properties of the CWMAD model are presented under the assumption of the specified restriction of the CWMF, which considers that the observation y_j is contained in the independent variable \mathbf{x}_j on the observed component.

The paper is organized in several sections and a subsection, where in Section 2 the Weighted Median of Data is presented. In Section 3, the approximation function $F: \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ is presented and analyzed. Afterwards, in Section 4, the CWMAD problem is defined, where it is shown that the minimization of the Δ can be conducted on the finite set. Then, in Subsection 4.1, the properties of Δ are presented, assuming the specified restriction of the CWMF. In Section 5, the numerical examples are given in order to present the performance of the CWMAD problem, where unequal dimensions of the independent variables are also observed. Finally, in Section 6, the conclusion is given.

2. The Weighted Median of Data

Let us denote a data vector $\mathbf{z} = (z_1, \dots, z_m) \in \mathbb{R}^m$, $m \in \mathbb{N}$, and a positive vector of weights $\mathbf{w} = (w_1, \dots, w_m) \in \mathbb{R}_+^m$. In this situation, we may define the function $f: \mathbb{R} \rightarrow \mathbb{R}$ as

$$f(u) = \sum_{i=1}^m w_i |z_i - u|, \quad (1)$$

which in that sense presents the LAD problem. Considering the minimization of the function f , which implies that the LAD regression can be observed as a problem of determining the appropriate global minimum of f . Furthermore, the convexity properties of the function f , can also be taken into consideration of the observed problem [4, 5]. In that sense, we may conclude that the global minimum of f always exists, which indicates the existence of the LAD solution [10], and thus it can be written that

$$\text{med}(\mathbf{w}, \mathbf{z}) = \underset{u \in \mathbb{R}}{\text{argmin}} f(u),$$

where $\text{med}(\mathbf{w}, \mathbf{z})$ is called the Weighted Median of Data [4, 8]. In general, the global minimum of f is not always unique, what confirms the well-known LAD regression properties of a possibility of a multiple solution existence.

The next theorem presents the problem of determining the weighted median of data $\text{med}(\mathbf{w}, \mathbf{z})$, which involves the problem of finding the global minimum of the function f , which in general is not unique.

Theorem 1. Let $\mathbf{z} \in \mathbb{R}^m$, $m \in \mathbb{N}$, be a data vector with a corresponding weight vector $\mathbf{w} \in \mathbb{R}_+^m$, and let $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(m)}$ be observations in an ascending order. Then it holds for the defined set

$$\mathcal{L} = \left\{ \ell: \sum_{i=1}^{\ell} w_{(i)} \leq \frac{W}{2} \right\}, \quad \ell \in \{1, \dots, m\}, \quad W = \sum_{i=1}^m w_i, \tag{2}$$

that:

- (a) if $\mathcal{L} = \emptyset$, then $\text{med}(\mathbf{w}, \mathbf{z}) = z_{(1)}$;
- (b) if $\mathcal{L} \neq \emptyset$, then for $\nu = \max \mathcal{L}$, it follows that:

(i) if $\sum_{i=1}^{\nu} w_{(i)} < \frac{W}{2}$, then $\text{med}(\mathbf{w}, \mathbf{z}) = z_{(\nu+1)}$;

(ii) if $\sum_{i=1}^{\nu} w_{(i)} = \frac{W}{2}$, then $\text{med}(\mathbf{w}, \mathbf{z}) = (1 - \lambda)z_{(\nu)} + \lambda z_{(\nu+1)}$, $\lambda \in [0, 1]$.

Proof. The function $f: \mathbb{R} \rightarrow \mathbb{R}$, defined by (1), is a piecewise linear function, so let us denote slopes κ_{ℓ} , $\ell \in \{0, \dots, m\}$, which correspond to each interval

$$\langle -\infty, z_{(1)} \rangle, \langle z_{(1)}, z_{(2)} \rangle, \dots, \langle z_{(m-1)}, z_{(m)} \rangle, \langle z_{(m)}, \infty \rangle.$$

In this situation, it follows that

$$\kappa_0 = -W, \quad \kappa_m = W, \tag{3}$$

and for $\ell \in \{1, \dots, m - 1\}$, it follows that

$$\kappa_{\ell} = 2 \sum_{i=1}^{\ell} w_{(i)} - W = \kappa_{\ell-1} + 2w_{(\ell)}. \tag{4}$$

- (a) Let us consider the case when $\mathcal{L} = \emptyset$. Then it holds that

$$\sum_{i=1}^{\ell} w_{(i)} > \frac{W}{2}, \quad \forall \ell \in \{1, \dots, m\}.$$

By using (3) and (4), we may conclude that $\kappa_0 < 0 < \kappa_{\ell}$, $\forall \ell \in \{1, \dots, m\}$, which implies that f is decreasing at $\langle -\infty, z_{(1)} \rangle$, and increasing at $\langle z_{(1)}, \infty \rangle$. Thereby, we may conclude that f reaches its global minimum at $z_{(1)}$, i.e. $\text{med}(\mathbf{w}, \mathbf{z}) = z_{(1)}$.

- (b) Let us consider the case when $\mathcal{L} \neq \emptyset$. This means that there exists $\nu = \max \mathcal{L}$. By using (4), we may conclude that $\kappa_{\nu} \leq 0$, while $\kappa_{\ell} > 0$, $\forall \ell > \nu$.

- (i) Let us consider the case when $\kappa_\nu < 0$. This means that f decreases at $\langle -\infty, z_{(\nu+1)} \rangle$, and increases at $\langle z_{(\nu+1)}, \infty \rangle$. Therefore, we may conclude that f reaches its global minimum at $z_{(\nu+1)}$, i.e. $\text{med}(\mathbf{w}, \mathbf{z}) = z_{(\nu+1)}$.
- (ii) Let us consider the case when $\kappa_\nu = 0$. This implies that f is decreasing at $\langle -\infty, z_{(\nu)} \rangle$, is constant on $[z_{(\nu)}, z_{(\nu+1)}]$, and is increasing at $\langle z_{(\nu+1)}, \infty \rangle$. Thereby, this implies that f reaches its global minimum at interval $[z_{(\nu)}, z_{(\nu+1)}]$, i.e. $\text{med}(\mathbf{w}, \mathbf{z}) = (1 - \lambda)z_{(\nu)} + \lambda z_{(\nu+1)}$, $\lambda \in [0, 1]$.

Remark 1. *If the problem of the weighted median of data is observed for weights, which are all set to one, i.e. $w_1 = \dots = w_m = 1$, then the global minimum of the corresponding function f is called the Median of Data, and is denoted as $\text{med}(\mathbf{z})$.*

3. The Component Weighted Median Function

Let us denote the component weighted model vector

$$\mathbf{w}_k = (1, \dots, w, \dots, 1), \quad w > 0,$$

where k -th component is observed as a variable, while all other components are set to one. In that sense, the Component Weighted Median Function (CWMF) $F: \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ is defined as

$$F(\mathbf{x}; w) = \text{med}(\mathbf{w}_k, \mathbf{x}).$$

The performance of the constructed function is studied through Theorem 1 by observing the parameter $w > 0$, and thus the CWMF restriction $F|_{\mathcal{D}}$ will be observed, where $\mathcal{D} = \mathbf{z} \times \mathbb{R}_+$, $\mathbf{z} \in \mathbb{R}^m$. In order to carry out some properties of $F|_{\mathcal{D}}$, a position $p \in \mathcal{P}$, $\mathcal{P} = \{q: z_{(q)} = z_k\}$, $q \in \{1, \dots, m\}$, of k -th component in ordered observation will be taken into consideration. Considering that, the two cases can be carried out: (I) a weight w is included into the sum; (II) a weight w is not included into the sum, which determines the set \mathcal{L} from Theorem 1 defined by (2).

- (I) The first case considered a situation when $w = w_{(p)}$, $p \in \mathcal{P}$, is included into the sum, i.e. $p \leq \ell$ and thus

$$\sum_{i=1}^{\ell} w_{(i)} = \ell - 1 + w.$$

Now it can be written that

$$\mathcal{L} = \{\ell: \ell \leq \ell^-(w)\}, \quad \ell^-(w) = \frac{m + 1 - w}{2}.$$

- (II) The second case is when $p \geq \ell + 1$, $p \in \mathcal{P}$, which means that a weight $w = w_{(p)}$ is not included into the sum, i.e.

$$\sum_{i=1}^{\ell} w_{(i)} = \ell.$$

In this situation

$$\mathcal{L} = \{\ell: \ell \leq \ell^+(w)\}, \quad \ell^+(w) = \frac{m-1+w}{2}.$$

The next theorem presents that the restricted CWMF is a piecewise constant function, and thereby the regions of constant values are investigated.

Theorem 2. *The function $F|_{\mathcal{D}}: \mathcal{D} \rightarrow \mathbb{R}$ is a piecewise constant function with respect to a parameter $w > 0$.*

Proof.

(I) Let us consider the first case. Following Theorem 1, it can be concluded that

$$\{z_{(p)}, \dots, z_{(\lceil \frac{m+1}{2} \rceil)}\} \subseteq \{F(\mathbf{z}, w): p \leq \ell\} \subseteq [z_{(p)}, z_{(\lceil \frac{m+1}{2} \rceil)}].$$

Considering the inequality

$$\nu_i^- < \ell^-(w) < \nu_i^- + 1, \quad z_{(\nu_i^-+1)} \in \{z_{(p)}, \dots, z_{(\lceil \frac{m+1}{2} \rceil)}\}, \tag{5}$$

it holds that $\nu_i^- = \max \mathcal{L}$, $w \in I_i$, where an interval I_i presents the solution of the observed inequality (5) for the corresponding ν_i^- . This implies, by the statment (b)-(i) of Theorem 1, that $F(\mathbf{z}; w) = z_{(\nu_i^-+1)}$, $w \in I_i$. In this situation, let us define the set Ψ^- in a descending order as

$$\Psi^- = \{\nu_i^-: \nu_1^- > \dots > \nu_s^-\}, \quad z_{(\nu_i^-+1)} \in \{z_{(p)}, \dots, z_{(\lceil \frac{m+1}{2} \rceil)}\}.$$

Then, the observed inequality (5) generates the intervals I_i for each ν_i^- :

$$I_1 = \langle 0, w_1^+ \rangle, \dots, I_i = \langle w_i^-, w_i^+ \rangle, \dots, I_s = \langle w_s^-, +\infty \rangle,$$

where $w_i^- = m - 1 - 2\nu_i^-$, and $w_i^+ = m + 1 - 2\nu_i^-$. The situation when the equality is observed, i.e.

$$\nu_i^- = \ell^-(w),$$

which corresponds to the case (b)-(ii) of Theorem 1. In this situation, it holds that $F(\mathbf{z}; w_i^+) = (1 - \lambda)z_{(\nu_i^-)} + \lambda z_{(\nu_i^-+1)}$, $\lambda \in [0, 1]$, where $\nu_i^- \in \Psi^- \setminus \{\nu_s^-\}$.

(II) Let us now consider the second case. Analogously, it can be concluded that

$$\{z_{(\lfloor \frac{m+1}{2} \rfloor)}, \dots, z_{(p-1)}\} \subseteq \{F(\mathbf{z}, w): p \geq \ell + 1\} \subseteq [z_{(\lfloor \frac{m+1}{2} \rfloor)}, z_{(p)}].$$

Then, the inequality

$$\nu_i^+ < \ell^+(w) < \nu_i^+ + 1, \quad z_{(\nu_i^++1)} \in \{z_{(\lfloor \frac{m+1}{2} \rfloor)}, \dots, z_{(p-1)}\},$$

is observed in order to indicate when $F(\mathbf{z}; w) = z_{(\nu_i^++1)}$. Again, we defined the set Ψ^+ , but now in an ascending order, i.e.

$$\Psi^+ = \{\nu_i^+ : \nu_1^+ < \dots < \nu_{s-1}^+\}, \quad z_{(\nu_i^++1)} \in \{z_{(\lfloor \frac{m+1}{2} \rfloor)}, \dots, z_{(p-1)}\}.$$

The results are obtained at intervals

$$I_1 = \langle 0, w_1^+ \rangle, \dots, I_i = \langle w_i^-, w_i^+ \rangle, \dots, I_{s-1} = \langle w_{s-1}^-, w_{s-1}^+ \rangle,$$

where $w_i^- = -m + 1 + 2\nu_i^+$, and $w_i^+ = -m + 3 + 2\nu_i^+$. The situation when the equality

$$\nu_i^+ = \ell^+(w)$$

is observed, generates that $F(\mathbf{z}; w_i^-) = (1 - \lambda)z_{(\nu_i^+)} + \lambda z_{(\nu_i^++1)}$, $\lambda \in [0, 1]$, where $\nu_i^+ \in \Psi^+ \setminus \{\nu_1^+\}$.

Remark 2. *The second case (II) always converts to the first case (I) when $w > |m + 1 - 2p|$. This is the situation when $w = w_{(p)}$ starts to be included into the sum (3), which immediately generates the interval I_s for the second case (II) in Theorem 2, i.e.*

$$I_s = \langle w_s^-, +\infty \rangle,$$

so that $F(\mathbf{z}; w) = z_{(\nu_s^++1)} = z_{(p)}$, $w \in I_s$, and singleton case when

$$F(\mathbf{z}; w_s^-) = (1 - \lambda)z_{(\nu_s^+)} + \lambda z_{(\nu_s^++1)},$$

where $w_s^- = w_{s-1}^+ = |m + 1 - 2p|$, and $\nu_s^+ = p - 1$.

Remark 3. *The case when the restricted CWMF is constant, i.e. $\text{Ima } F|_{\mathcal{D}} = \{z_k\}$ is singleton, implies that $F(\mathbf{z}; w) = z_k$, $w \in I_1 = \mathbb{R}_+$.*

As it is shown, the restricted CWMF is a piecewise constant function with finitely many pieces. Therefore, according to Theorem 2, Remark 2 and 3, the restricted CWMF can be written as a finite linear combination

$$F(\mathbf{z}; w) = \sum_{t=2}^{2s} \alpha_t^{\text{sgn}} \chi_{A_t}(w), \quad \text{sgn} = \begin{cases} +, & p > \frac{m+1}{2}; \\ -, & p \leq \frac{m+1}{2}, \end{cases} \quad (6)$$

where

$$\alpha_t^\pm = \begin{cases} z_{(\nu_i^\pm+1)}, & t = 2i; \\ (1 - \lambda)z_{(\nu_i^\pm)} + \lambda z_{(\nu_i^\pm+1)}, & t = 2i \pm 1; \end{cases} \quad A_t = \begin{cases} I_i, & t = 2i; \\ \{w_i^\mp\}, & t = 2i \pm 1, \end{cases}$$

and χ_A presents the indicator function, which is defined as

$$\chi_A(w) = \begin{cases} 1, & w \in A; \\ 0, & w \notin A. \end{cases}$$

The next figure presents the performance of the restricted CWMF, where the dashed line presents the situation for the first case (I), and the solid line for the second case (II). According to Remark 2, it can be seen that the first case (I) also starts to appear for the second case (II) when $w > |m+1-2p|$. The Figure 1(a) presents $F|_{\mathcal{D}}$, $\mathcal{D} = \mathbf{z} \times \mathbb{R}_+$, $\mathbf{z} \in \mathbb{R}^m$, ($m = 8$), when $\lambda = 0$ is applied, where the red marked graph denotes the situation when $p > \frac{m+1}{2}$ ($p = 8$), and blue when $p \leq \frac{m+1}{2}$ ($p = 1$). Analogously, Figure 1(b) presents $F|_{\mathcal{D}}$, $\mathcal{D} = \mathbf{z} \times \mathbb{R}_+$, $\mathbf{z} \in \mathbb{R}^m$, ($m = 9$), when $\lambda = 1$ is applied. It can be seen that in this situation, the regions of the constant values of $F|_{\mathcal{D}}$ shift, what is caused by an odd number of observations.

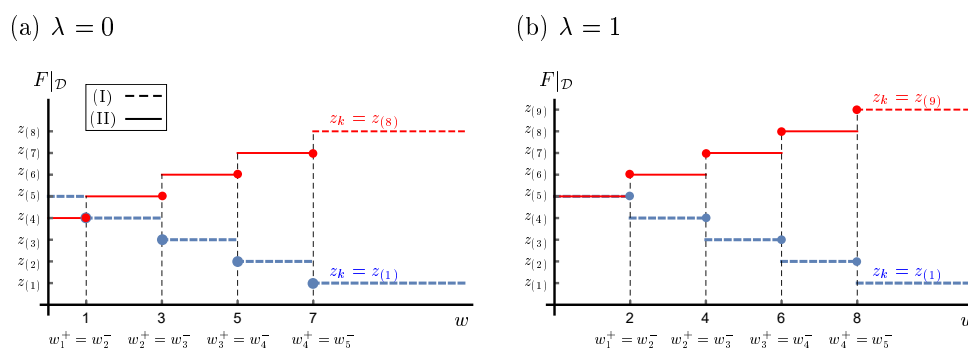


Figure 1: The restricted CWMF

Remark 4. *The restricted CWMF follows its monotonicity according to the position of the observed component in ordered observation, i.e.*

$$F(\mathbf{z}; w) = \begin{cases} \text{monotonically increasing,} & p > \frac{m+1}{2}; \\ \text{monotonically decreasing,} & p \leq \frac{m+1}{2}. \end{cases}$$

The next corollary considers the restricted CWMF when $\lambda \in \{0, 1\}$ is taken into consideration.

Corollary 1. *It holds for $F|_{\mathcal{D}}: \mathcal{D} \rightarrow \mathbb{R}$ that:*

- (a) when $\lambda = 0$, then:
 - (i) if $F|_{\mathcal{D}}$ decreases, then it is right continuous;
 - (ii) if $F|_{\mathcal{D}}$ increases, then it is left continuous.
- (b) when $\lambda = 1$, then:
 - (i) if $F|_{\mathcal{D}}$ decreases, then it is left continuous;
 - (ii) if $F|_{\mathcal{D}}$ increases, then it is right continuous;

Proof. Considering (6) for $\lambda \in \{0, 1\}$, the restricted CWMF can be written as

$$F(\mathbf{z}; w) = \sum_{t=1}^s \alpha_t^{\text{sgn}} \chi_{A_t}(w),$$

where $\alpha_t^\pm = z_{(v_t^\pm+1)}$, and A_t can be presented as left-open intervals:

$$A_1 = \langle 0, w_1^+ \rangle, \dots, A_t = \langle w_t^-, w_t^+ \rangle, \dots, A_s = \langle w_s^-, +\infty \rangle,$$

or right-open intervals:

$$A_1 = \langle 0, w_1^+ \rangle, \dots, A_t = [w_t^-, w_t^+), \dots, A_s = [w_s^-, +\infty),$$

and thus the statements of the corollary are proven.

Remark 5. *It follows that $F(\mathbf{z}; w) = z_{(p)} = z_k$, when $w > |m + 1 - 2p^*|$, where*

$$p^* = \begin{cases} \min\{p^-, \frac{m+1}{2}\}, & p > \frac{m+1}{2}; \\ \max\{p^+, \frac{m+1}{2}\}, & p \leq \frac{m+1}{2}, \end{cases}$$

such that

$$p^+ = \max \mathcal{P}, \quad p^- = \min \mathcal{P}, \quad \mathcal{P} = \{q: z_{(q)} = z_k\}.$$

This means that the restricted CWMF converges to the k -th component of $\mathbf{z} \in \mathbb{R}^m$, i.e.

$$\lim_{w \rightarrow +\infty} F(\mathbf{z}; w) = z_k.$$

4. The LAD regression with the CWMF

In this section we present the LAD regression model, which considers the CWMF as an approximation function, whose value is used to predict the outcome of a dependent variable. It is well known that the L_1 norm has been widely used to make robust models, which is useful in preventing model misspecifications, which are often caused by outliers [7]. The problem is considered as the L_1 norm error model function

$$\Delta(w) = \sum_{j=1}^n |y_j - F(\mathbf{x}_j; w)|,$$

where $Y = \{y_j \in \mathbb{R}: j \in \{1, \dots, n\}\}$ presents the dependent variables, $X = \{\mathbf{x}_j \in \mathbb{R}^m: j \in \{1, \dots, n\}\}$ the independent variables, and $w > 0$ the model parameter. In this situation, the robust regression model, which considers the Component Weighted Median Absolute Deviations (CWMAD), is constructed.

Considering the restricted CWMF as a piecewise constant function with finitely many pieces (6), the minimization problem of Δ can be conducted on a finite set

$$\mathcal{A} = \{a_t: a_t \in A_t\}, \quad t \in \{2, 3, \dots, 2s\}, \quad s = \max_{j \in \{1, \dots, n\}} s_j, \tag{7}$$

where s_j denotes a number of intervals of $F|_{\mathcal{D}_j}$, $\mathcal{D}_j = \mathbf{x}_j \times \mathbb{R}_+$. In this case, it follows that

$$\min_{w>0} \Delta(w) = \min_{a \in \mathcal{A}} \Delta(a),$$

which implies that the global minimum of Δ satisfies that

$$\Delta(w^*) = \min_{w>0} \Delta(w), \quad w^* \in A^* = \bigcup_{t \in \mathcal{T}} A_t,$$

where $\mathcal{T} = \{t: \Delta(a_t) = \min_{a \in \mathcal{A}} \Delta(a)\}$.

4.1. The CWMAD properties

In this subsection, the properties of the CWMAD are presented with respect to the specified restriction of the CWMF. The restriction $F|_{\overline{\mathcal{D}}_j}$ is constructed in such a way that the k -th component of an independent variable is equal to a dependent variable for each observation, i.e. $\overline{\mathcal{D}}_j = \mathbf{x}_j \times \mathbb{R}_+$, $x_k^{(j)} = y_j$, where $\mathbf{x}_j = (x_1^{(j)}, \dots, x_m^{(j)}) \in \mathbb{R}^m$. Thereby, the residual functions $r_j: \mathbb{R}_+ \rightarrow \mathbb{R}$ of the CWMAD

$$r_j(w) = y_j - F|_{\overline{\mathcal{D}}_j}(\mathbf{x}; w),$$

are studied in order to determine the global minimum of Δ , i.e. the optimal model parameter.

In the next lemma, the properties of the residuals r_j are briefly listed.

Lemma 1. *It holds for the residual functions $r_j: \mathbb{R}_+ \rightarrow \mathbb{R}$ that:*

- (a) r_j is a monotonically piecewise constant function;
- (b) $r_j(w^*) = \min_{w>0} |r_j(w)|$, $w^* \in A_j^*$, where

$$A_j^* = \begin{cases} [w_j^*, +\infty), & \text{if } r_j \text{ is right continuous;} \\ \langle w_j^*, +\infty), & \text{if } r_j \text{ is left continuous or constant,} \end{cases}$$

such that $w_j^* = |m + 1 - 2p_j^*|$.

Proof.

- (a) According to (6), $F|_{\overline{\mathcal{D}}_j}$ is a monotonically piecewise constant function, which according to Remark 5, reaches k -th component $x_k^{(j)} = y_j$ at its last piece. Considering that, we may conclude that the residuals are also monotonically piecewise constant functions, which can be written as

$$r_j(w) = \sum_{t=2}^{2s_j} \alpha_t^{(j)} \chi_{A_t}(w), \quad \alpha_t^{(j)} = y_j - \alpha_t^{\text{sgn}_j}, \quad \text{sgn}_j = \begin{cases} +, & p_j > \frac{m+1}{2}; \\ -, & p_j \leq \frac{m+1}{2}, \end{cases}$$

where s_j presents an interval number of $F|_{\overline{\mathcal{D}}_j}$, $p_j \in \mathcal{P}_j = \{q: x_{(q)}^{(j)} = x_k^{(j)}\}$, and

$$\alpha_t^{(j)} = \begin{cases} y_j - x_{(\nu_i^{\pm}+1)}^{(j)}, & t = 2i; \\ y_j - F(\mathbf{x}_j; w_i^{\mp}), & t = 2i \pm 1; \\ 0, & t \geq 2s_j. \end{cases}$$

- (b) Considering the previous statement (a), the residuals r_j reach zero value at their least piece. Thereby, the minimization problem

$$\min_{w>0} |r_j(w)|,$$

reaches its global minimum at a right-unbounded interval, which we denote as A_j^* . Considering $F|_{\overline{\mathcal{D}}_j}$ when it fulfills the statement (a)-(i), or (b)-(ii) of Corollary 1, we may conclude that r_j is right continuous, and thus A_j^* is closed from the left side by $w_j^* = |m - 1 + 2p_j^*|$, where according to Remark 5, it follows that

$$p_j^* = \begin{cases} \min\{p_j^-, \frac{m+1}{2}\}, & p_j > \frac{m+1}{2}; \\ \max\{p_j^+, \frac{m+1}{2}\}, & p_j \leq \frac{m+1}{2}, \end{cases}$$

so that

$$p_j^+ = \max \mathcal{P}_j, \quad p_j^- = \min \mathcal{P}_j, \quad \mathcal{P}_j = \{q: x_{(q)}^{(j)} = x_k^{(j)}\}.$$

Otherwise, if $F|_{\overline{\mathcal{D}}_j}$ is constant, or $\lambda \in (0, 1)$, then A_j^* is a right-unbounded interval which is open from the left by w_j^* .

Theorem 3. *It holds for $\Delta: \mathbb{R}_+ \rightarrow \mathbb{R}_+$, defined as*

$$\Delta(w) = \sum_{j=1}^n |r_j(w)|,$$

that:

- (a) Δ is a monotonically decreasing piecewise constant function;
- (b) $\Delta(w^*) = \min_{w \in \mathbb{R}} \Delta(w)$, $w^* \in A^* = \bigcap_{j=1}^n A_j^*$.

Proof.

- (a) Considering Lemma 1 (a), it follows that $|r_j|$ is a monotonically decreasing piecewise constant function, which converges to zero, and thus we can write that

$$\Delta(w) = \sum_{t=2}^{2s} \alpha_t \chi_{A_t}(w), \quad s = \max_{j \in \{1, \dots, n\}} s_j, \quad \alpha_t = \sum_{j=1}^n |\alpha_t^{(j)}|.$$

Knowing that the sum of the monotonically decreasing functions is also decreasing, we may conclude that the statement (a) is proven.

- (b) Considering the Lema 1 (b), we can conclude that Δ reaches the global minimum when all residuals r_j are zero. Because all A_j^* are right-unbounded intervals, it holds that the global minimum of the observed problem is contained in the intersections of all A_j^* .

5. Numerical examples

In this section, the numerical examples for the generalized CWMAD problem are presented, where the independent variables are considered to be with unequal dimensions, i.e. $X = \{\mathbf{x}_j \in \mathbb{R}^{m_j} : j \in \{1, \dots, n\}\}$, $m_j \in \mathbb{N}$, and thus with the different component weighted model vector of the CWMF. In this situation, the L_1 norm error model function is defined as

$$\Delta(w) = \sum_{j=1}^n |y_j - F_j(\mathbf{x}_j; w)|,$$

where the functions $F_j: \mathbb{R}^{m_j+1} \rightarrow \mathbb{R}$ present the CWMF defined as

$$F_j(\mathbf{x}; w) = \text{med}(\mathbf{w}_{k_j}, \mathbf{x}), \quad 1 \leq k_j \leq m_j.$$

In the next table, the numerical examples are presented in order to show the performance of Δ ($\lambda = 0$ is observed), together with the results of the generalized CWMAD problem. The optimal results are obtained by taking into account the appropriate finite set defined by (7). In that sense, the finite set must be constructed to present the regions of the corresponding model function Δ , which contains the finite pieces of the constant values. For that purpose, the approximating finite set $\tilde{\mathcal{A}}$ ($\supseteq \mathcal{A}$) is constructed, which presents the regions of the model function Δ for each numerical example presented in Table 1, i.e.

$$\tilde{\mathcal{A}} = \{0.5i : i \in \{1, \dots, 2M\}\}, \quad M = \max_{j \in \{1, \dots, 10\}} m_j,$$

and thus

$$\min_{a \in \tilde{\mathcal{A}}} \Delta(w) = \min_{w > 0} \Delta(w).$$

The next figure presents the performance of Δ , where Figures 2(a),(b), and (c) correspond to the numerical examples (a), (b), and (c) from Table 1. It can be seen in Figure 2(b) that the example (b) from Table 1, is constructed according to the specified restriction, where the dependent variable is contained on the observed component for each independent variable (see Subsection 4.1). According to Theorem 3, it follows that the observed model function Δ is monotonically decreasing, and attains its global minimum at its least piece.

Table 1: Numerical examples

No.	(a)				(b)				(c)						
j	x_j			y_j	k_j	x_j			y_j	k_j	x_j			y_j	k_j
1	(4, 4, 4, 8, 2, 5, 5, 9)			1	4	(3, 0, 8, 5, 9, 0, 6, 7, 2)			0	2	(6, 5, 0, 9)			0	4
2	(5, 8, 4, 5, 2, 2, 6, 8)			4	3	(2, 2, 6, 0, 7, 3, 9, 4, 1)			3	6	(7, 6, 9, 7, 3, 9, 7)			6	5
3	(3, 6, 3, 7, 0, 3, 6, 0)			3	8	(3, 8, 9, 8, 8, 6, 9, 8, 4)			8	8	(9, 4, 3)			5	2
4	(4, 5, 8, 4, 1, 6, 1, 4)			8	2	(4, 4, 8, 4, 5, 3, 3, 7, 9)			8	3	(4, 9)			3	1
5	(2, 2, 5, 8, 5, 9, 4, 6)			4	4	(0, 4, 3, 8, 2, 1, 5, 7, 3)			4	2	(6, 7, 1, 6)			0	4
6	(7, 1, 4, 7, 5, 0, 8, 8)			6	2	(7, 7, 8, 5, 8, 0, 4, 6, 7)			6	8	(9, 7, 4, 5, 3)			1	4
7	(2, 8, 1, 5, 8, 0, 0, 8)			5	3	(1, 7, 0, 1, 0, 4, 0, 5, 7)			1	4	(2, 5, 2, 7, 0, 0, 2)			4	6
8	(5, 1, 6, 7, 5, 5, 6, 9)			2	7	(9, 7, 6, 0, 9, 6, 7, 0, 6)			7	7	(0, 0, 7, 4, 2, 6)			0	2
9	(7, 8, 8, 5, 5, 1, 2, 9)			4	6	(2, 7, 6, 7, 0, 1, 7, 2, 6)			7	7	(4, 2)			7	2
10	(6, 6, 7, 8, 7, 9, 9, 4)			6	8	(6, 1, 0, 8, 2, 3, 0, 6, 4)			2	5	(0, 4, 1, 6)			9	3
$\Delta(w^*)$	26				0				34						
A^*	{5, 7}				{6, +∞}				{3}						

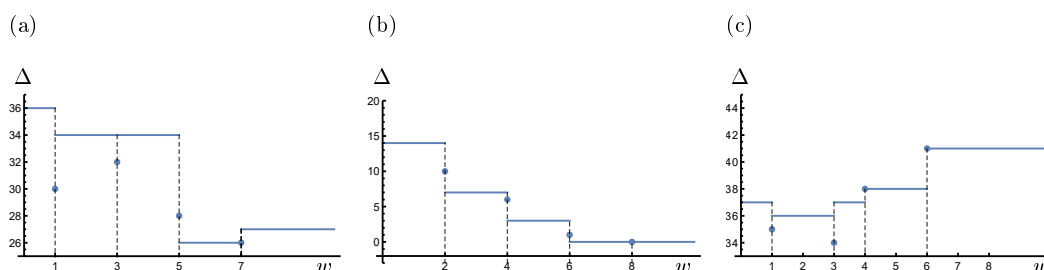


Figure 2: The model function Δ

6. Conclusion

It is shown that the designed CWMAD model estimates the optimal model parameter at the regions of constant values of the restricted approximation CWMF. Furthermore, the restricted CWMF is studied and analyzed, where the assumption of the specified restriction is considered in order to examine the CWMAD properties. Considering the numerical experiments for the observed problem, the generalized CWMAD problem is considered, where it is shown that the optimal parameter model is detected by designing the correct discrete optimization model.

Acknowledgements

This work has been fully supported by Mechanical Engineering Faculty in Slavonski Brod under the internal research project Generalization of Arithmetic Means and Their Applications.

We are grateful to Željka Rosandić who has linguistically corrected and refined this paper.

References

- [1] P Bloomfield and W Steiger. *Least Absolute Deviations: Theory, Applications and Algorithms*. Birkhauser, Boston, 1983.
- [2] V Novoselac and Z Pavić. Adaptive center weighted median filter. In A Sedmak, editor, *Proceedings of 7th International Scientific and Expert Conference of the International TEAM Society.*, pages 286–291, Belgrade, 2015. Faculty of Mechanical Engineering, University of Belgrade.
- [3] V Novoselac and Z Pavić. Cluster detection in noisy environment by using the modified EM algorithm. *Croatian Operational Research Review*, 9(2):223–234, 2018.
- [4] V Novoselac and Z Pavić. Optimal Solution Properties of an Overdetermined System of Linear Equations. *European Journal of Pure and Applied Mathematics*, 12(4):1360–1370, 2019.
- [5] M R Osborne. *Finite Algorithms in Optimization and Data Analysis*. John Wiley, Australian National University, Camberra, 1985.
- [6] Young Woong Park. Optimization for L_1 -Norm Error Fitting via Data Aggregation. *Inform Journal on Computing*, 2020.
- [7] P J Rousseeuw and A M Leroy. *Robust Regression and Outlier Detection*. Wiley, New York, 2003.
- [8] I Vazler K Sabo and R Scitovski. Weighted median of the data in solving least absolute deviations problems. *Communications in Statistics-Theory and Methods*, 41(8):1455–1465, 2012.
- [9] K Sabo and R Scitovski. The best least absolute deviations line-properties and two efficient methods for its derivation. *ANZIAM Journal*, 50(2):185–198, 2008.
- [10] G A Watson. *Approximation Theory and Numerical Methods*. John Wiley & Sons, New York, 1980.