Hacettepe Journal of Mathematics & Statistics

Hacet. J. Math. Stat.
Volume 54 (2) (2025), 762 – 778
DOI : 10.15672/hujms.1461429

Research Article

Robust regression analysis using the weighted median model for improved denoising of MR data in image processing

Vedran Novoselac*[1], Anita Katić[2]

[1] *Mechanical Engineering Faculty in Slavonski Brod, University of Slavonski Brod, Trg Ivane Brlić Mažuranić 2, HR-35 000 Slavonski Brod, Croatia*

[2] *The Department of Social Sciences and Humanities, University of Slavonski Brod, Trg Ivane Brlić Mažuranić 2, HR-35 000 Slavonski Brod, Croatia*

Abstract

This paper presents a comprehensive analysis of a multidimensional regression model using the weighted median as the regression function. The model is formulated as an optimization problem within the framework of the L_1-norm error fitting approach, exhibiting robustness to outliers, a critical advantage in various applications where data might be contaminated by extreme values. The core of the investigation focuses on the regression and objective functions of the proposed model. A detailed mathematical study reveals that the optimization problem inherent in the model can be effectively discretized, leading to computationally tractable solutions. The study's findings are further validated through a rigorous exploration of the model's application in the context of image denoising, a significant problem in image processing. Specifically, the model addresses the challenging task of impulse noise removal in Magnetic Resonance images. By integrating the proposed model into well-established adaptive denoising methods, this work demonstrates that significant improvements in image quality reconstruction and noise suppression are easily achievable. The results highlight the model's efficacy in balancing the competing demands of preserving essential image features while effectively reducing noise artifacts. This research offers a novel approach for robust regression analysis and provides a robust tool for image denoising, particularly in scenarios involving impulse noise. The mathematical underpinnings, along with the demonstrated practical application, contribute significantly to the field of robust statistical modeling and image processing.

Mathematics Subject Classification (2020). 62J02, 90C26, 46N10, 94A08

Keywords. functional analysis, image processing, optimization, regression analysis, robust regression

*Corresponding Author.
Email addresses: Vedran.Novoselac@unisb.hr (V. Novoselac), Anita.Katic@unisb.hr (A. Katić)
Received: 29.03.2024; Accepted: 12.03.2025

1. Introduction

In statistical learning models, data fitting is a crucial task in analysis. This process involves fitting a model to observed data and evaluating the resulting errors to evaluate the quality of the fit. A widely used approach for data fitting is the method of least squares (LS problem), which utilizes the L_2-norm to measure fitting errors. Given that the L_2-norm is sensitive to outliers, it is often necessary to use the L_1-norm to achieve better robustness in model construction [1,9,17]. The L_1-norm approach, commonly referred to as the Least Absolute Deviation (LAD) method, is widely used for constructing robust models and is based on the assumption that errors follow a double exponential (Laplace) distribution. This concept is attributed to Josip Ruđer Bošković (1711-1787), a Croatian scientist (mathematician, physicist, astronomer and philosopher) born in Dubrovnik. The principle was also applied by the French mathematician and astronomer Marquis Pierre-Simon de Laplace (1749-1827) [2]. While many L_2-norm-based models can be expressed analytically, solutions for L_1-norm models are often more complex and non-trivial. As a result, L_1-norm error fitting models are receiving increasing attention in recent research [10,12,14,15,18,22].

The aim of this investigation is to optimize the model

$$\Delta(\lambda^*, w^*) = \min_{\substack{\lambda\in[0,1]\\ w>0}} \Delta(\lambda, w), \tag{1.1}$$

where

$$\Delta(\lambda, w) = \sum_{j=1}^{m} |y_j - F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j)| \tag{1.2}$$

denotes the objective function in the context of the L_1-norm error fitting approach, while the function $(\lambda, w) \mapsto F(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j)$ represents the regression function, where $\boldsymbol{x}_j = (x_1^{(j)}, \ldots, x_n^{(j)}) \in \mathbb{R}^n$ denotes the independent variables and $y_j \in \mathbb{R}$ represents the dependent variables. The regression function of the observed model is defined as the weighted median, which has widespread applications in many fields of applied science and statistics, where its robust property against outliers comes to the fore [11,18,22]. Thus, the regression function is defined as

$$F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j) = \mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}_j), \tag{1.3}$$

where $\boldsymbol{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n_+$ denotes the weight vector, with each weight corresponding to a specific element of $\boldsymbol{x}_j = (x_1^{(j)}, \ldots, x_n^{(j)}) \in \mathbb{R}^n$. The first variable $\lambda \in [0, 1]$ in the regression function (1.3) represents a parameter that determines the weighted median in cases where it is not unique, while the second variable corresponds to the specific weight $w = w_k > 0$ of the observed weighted vector. To resolve the optimization problem (1.1), this study investigates the regression function (1.3) and the objective function (1.2), demonstrating that the observed optimization problem can be presented discretely.

It is well known that the observed weighting problem is particularly relevant in the field of image processing for image denoising, where specific weights help achieve a balance between image quality reconstruction and noise suppression [3, 16]. To implement the proposed modifications to the advanced image denoising methods, a generalized objective model is constructed that accommodates flexible input dimensionalities and weighting positions. Consequently, this generalized modification is applied to various methods that exhibit flexible performance [5,8,20,21], as well as to methods that are static [7,19,23]. The experimental results demonstrate that the generalized model enhances standard methods, resulting in notable improvements in suppressing well-known impulse noise in Magnetic Resonance (MR) images.

The study is presented in several sections and subsections. In Section 2, the weighted median of the data and its properties are presented. In Section 3 the analysis of the observed regression model is performed, where the regression and the objective functions are studied and described in two subsections, Subsection 3.1 and Subsection 3.2, respectively. Afterward, the optimization process of the observed model (1.1) is described in Section 4, where it is shown that the optimization process can be presented discretely. In Section 5, a numerical example is presented in Subsection 5.1, where the observed problem is applied to estimate the expectations of normally distributed data in the presence of outliers. Afterward, Subsection 5.2 presents the application of the observed model for image denoising methods, demonstrating that the proposed modifications achieve significant results. And finally, in Section 6 the conclusion is given.

2. The weighted median

The weighted median of $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ with the corresponding weights $\boldsymbol{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n_+$ is defined as the estimate of the LAD problem that can be presented as the global minimum of the function

$$\begin{aligned} f(x; \boldsymbol{w}, \boldsymbol{x}) &= \sum_{i=1}^{n} w_i |x_i - x|; \\ &= \sum_{t=1}^{n+1} f_t(x)\ \mathbf{1}_{I_t}(x), \end{aligned} \tag{2.1}$$

where $f_t(x) = \kappa_t\, x + \ell_t$ present linear function defined on the particular interval I_t. In this situation

$$\mathbf{1}_{I_t}(x) = \begin{cases} 1, & x \in I_t; \\ 0, & x \notin I_t, \end{cases}$$

denotes the indicator function, where intervals I_t are defined as

$$I_1 = \langle -\infty, x_{\pi(1)}], \ldots, I_t = \langle x_{\pi(t-1)}, x_{\pi(t)}], \ldots, I_{n+1} = \langle x_{\pi(n)}, +\infty\rangle,$$

such that

$$x_{\pi(1)} \leq x_{\pi(2)} \leq \cdots \leq x_{\pi(n)},$$

denotes sorted elements of $\boldsymbol{x} \in \mathbb{R}^n$ in the ascending order that is obtained by some permutation $\pi \in S_n$ (S_n is permutation group of degree n). The function (2.1) is linear and convex in piecewise order, ensuring the existence of its global minimum [12, 13, 15]. The method for determining the weighted median is well known [4], and is based on the principle that its global minimum occurs at the point where the coefficient κ_t of f_t transitions from a negative to a positive value. Therefore, considering the previous discussions, the following theorem presents the determination of the weighted median.

Theorem 2.1. *Let $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ be a vector of values with corresponding weights $\boldsymbol{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n_+$, and let $\pi \in S_n$ be a permutation that sorts the elements of $\boldsymbol{x}$ in ascending order. Then for $s = \max \mathcal{T}$, where*

$$\mathcal{T} = \{t \mid 2\sum_{i=1}^{t-1} w_{\pi(i)} \leq \sum_{i=1}^{n} w_{\pi(i)}\}, \quad t \in \{1, \ldots, n\},$$

it holds that:

(i) $2\sum_{i=1}^{s-1} w_{\pi(i)} < \sum_{i=1}^{n} w_{\pi(i)} \implies \mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) = x_{\pi(s)}$;

(ii) $2\sum_{i=1}^{s-1} w_{\pi(i)} = \sum_{i=1}^{n} w_{\pi(i)} \implies \mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) = (1-\lambda) x_{\pi(s-1)} + \lambda x_{\pi(s)}$, $\lambda \in [0, 1]$.

Proof. Let us determine the coefficients κ_t of $f_t(x)$ in (2.1), i.e.

$$\kappa_t = 2\sum_{i=1}^{t-1} w_{\pi(i)} - \sum_{i=1}^{n} w_{\pi(i)}. \tag{2.2}$$

By insight in (2.2), it may be concluded that the coefficients κ_t are strictly increasing, i.e. it holds that

$$-\sum_{i=1}^{n} w_{\pi(i)} = \kappa_1 < \kappa_2 < \cdots < \kappa_{n+1} = \sum_{i=1}^{n} w_{\pi(i)}. \tag{2.3}$$

In order to prove (i), let us consider the situation when $\kappa_s < 0$. Then, according to the definition of the set $\mathcal{T}$ and the statement (2.3), it can be concluded that $\kappa_t < 0$, $\forall t \le s$, and $\kappa_t > 0$, $\forall t > s$. This means that function $f(x; \boldsymbol{w}, \boldsymbol{x})$ strictly decreases on interval $\langle -\infty, x_{\pi(s)} \rangle$, and strictly increases on interval $\langle x_{\pi(s)}, +\infty \rangle$. Thus, it can be concluded that the global minimum is reached at $x_{\pi(s)}$, i.e. $\mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) = x_{\pi(s)}$.

Now, let us consider the statment (ii) when $\kappa_s = 0$. Analogously as in the first case, it can be concluded that $\kappa_t < 0$, $\forall t \le s-1$, and $\kappa_t > 0$, $\forall t > s$. This means that the function $f(x; \boldsymbol{w}, \boldsymbol{x})$ strictly decreases in the interval $\langle -\infty, x_{\pi(s-1)} \rangle$, stagnates in the segment $[x_{\pi(s-1)}, x_{\pi(s)}]$, and increases strictly in the interval $\langle x_{\pi(s)}, +\infty \rangle$. In this situation, it can be concluded that the global minimum is reached on $[x_{\pi(s-1)}, x_{\pi(s)}]$, that is, $\mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) = (1-\lambda)x_{\pi(s-1)} + \lambda x_{\pi(s)}$, $\lambda \in [0, 1]$.

□

Remark 2.2. If equal weights are considered, that is, $w_1 = \cdots = w_n$, then the weighted median is called the median and is denoted as $\mathrm{med}_\lambda(\boldsymbol{x})$.

Corollary 2.3. *Let $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\boldsymbol{w} = (w_1, \dots, w_n) \in \mathbb{R}^n_+$ and $\pi \in S_n$ which sorts $\boldsymbol{x}$ in ascending order. Then, for some $k \in \{1, \dots, n\}$, the following implications hold:*

(i) $\delta_s a_{s+1} < \delta_s w_k < \delta_s a_s \implies \mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) = x_{\pi(s)}$;

(ii) $w_k = a_s \implies \mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) = (1-\lambda)x_{\pi(s-1)} + \lambda x_{\pi(s)}$;

(iii) $w_k > |a_l| = |a_{l+1}| \implies \mathrm{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) = x_{\pi(l)} = x_k$,

where $l = \pi^{-1}(k)$ and

$$a_t = \begin{cases} \displaystyle\sum_{i=1}^{t-1} w_{\pi(i)} - \sum_{\substack{i=t \\ i \ne l}}^{n} w_{\pi(i)}, & t \le l; \\ \displaystyle\sum_{i=t}^{n} w_{\pi(i)} - \sum_{\substack{i=1 \\ i \ne l}}^{t-1} w_{\pi(i)}, & t \ge l+1, \end{cases} \qquad \delta_t = \begin{cases} -1, & t \le l; \\ 1, & t \ge l+1. \end{cases} \tag{2.4}$$

Proof. Let us rewrite the set $\mathcal{T}$ from Theorem 2.1 in order to observe $w_k > 0$ as follows

$$\mathcal{T} = \{t \mid \delta_t \, w_k \le \delta_t \, a_t\}, \quad t \in \{1, \dots, n\},$$

where the coefficients a_t and Kronecker delta δ_t are defined by (2.4). Considering (2.4), it can be concluded that the coefficients a_t strictly increase when $t \le l$, i.e. it holds that

$$-\sum_{\substack{i=1 \\ i \ne l}}^{n} w_{\pi(i)} = a_1 < a_2 < \cdots < a_l, \tag{2.5}$$

while in situation when $t \geq l+1$ coefficients strictly decrease, i.e.

$$a_{l+1} > a_{l+2} > \cdots > a_{n+1} = -\sum_{\substack{i=1\\ i\neq l}}^{n} w_{\pi(i)}. \tag{2.6}$$

Finally, considering equations (2.4), (2.5), (2.6), and Theorem 2.1, the statements of the corollary are clearly established, illustrating the relationship between the weight w_k and the weighted median. $\square$

Proposition 2.4. *Let α, β, $\gamma \in \mathbb{R}$, such that $\alpha > 0$, $\beta \neq 0$, then it holds that*

$$\operatorname{med}_\lambda(\alpha \boldsymbol{w}, \beta \boldsymbol{x} + \gamma \boldsymbol{e}) = \beta \operatorname{med}_\lambda(\boldsymbol{w}, \boldsymbol{x}) + \gamma,$$

where $\boldsymbol{e} = (1, \ldots, 1) \in \mathbb{R}^n$.

Proof. Let us denote substitutions as $z = \beta x + \gamma$. It can be directly concluded from (2.1) that it holds that

$$f(z; \alpha \boldsymbol{w}, \beta \boldsymbol{x} + \gamma \boldsymbol{e}) = \alpha\, |\beta|\, f(x; \boldsymbol{w}, \boldsymbol{x}).$$

In this situation, the left side is at its minimum at $z^* = \operatorname{med}_\lambda(\alpha \boldsymbol{w}, \beta \boldsymbol{x} + \gamma \boldsymbol{e})$, while the right side attends at $x^* = \operatorname{med}_\lambda(\boldsymbol{w}, \boldsymbol{x})$. Finally, applying the substitution, we can conclude that $z^* = \beta x^* + \gamma$, which proves the proposition. $\square$

Example 2.5. The next figure presents $f(x; \boldsymbol{w}, \boldsymbol{x})$ for $\boldsymbol{x} = (1, 2, 3, 4, 5)$. It shows that the observed function always reaches its global minimum, which, as stated in (ii) of Theorem 2.1, is not always unique. In Figure 1(a), we see the situation when $\boldsymbol{w} = (1, 2, 3, 2, 1)$, which generates a unique global minimum reached at $x_{\pi(3)} = 3$. In contrast, Figure 1(b) considers $\boldsymbol{w} = (1, 1, 3, 4, 1)$, which generates a global minimum throughout the segment $[x_{\pi(3)}, x_{\pi(4)}] = [3, 4]$.

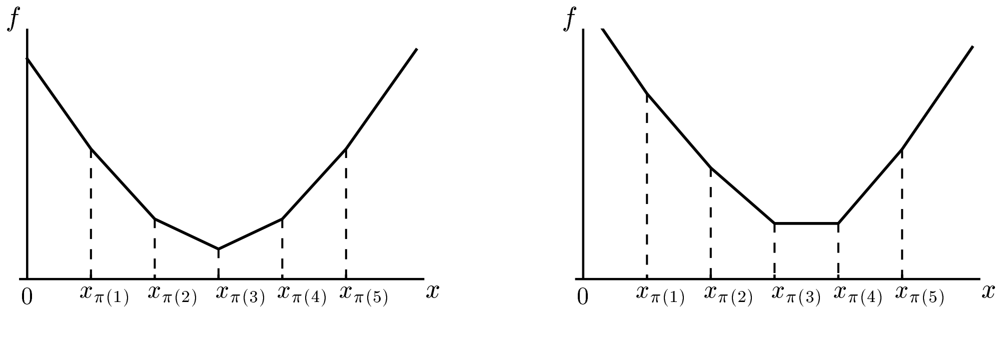


Figure 1. Graph of f

3. The component weighted median regression analysis

In this section, the analysis of the observed model (1.1) is conducted and is presented in two subsections. In the first subsection, the regression function (1.3) is studied, while in the second subsection the objective function (1.2) is examined.

3.1. The regression function analysis

In order to analyze the median-based regression function, the previous Section 2 is referenced to present and discuss some properties of the regression function.

Theorem 3.1. *Let $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, $\boldsymbol{w} = (w_1, \dots, w_n) \in \mathbb{R}^n_+$ and $\pi \in S_n$ which sorts $\boldsymbol{x}$ in ascending order. Then, for some $k \in \{1, \dots, n\}$ and $l = \pi^{-1}(k)$, it holds that*

$$F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}) = \sum_{t=1}^{2n-1} \alpha_t(\lambda)\ \mathbf{1}_{A_t}(w),$$

where

$$\alpha_t(\lambda) = \begin{cases} x_{\pi(s)}, & t = 2s-1; \\ (1-\lambda)x_{\pi(s-1)} + \lambda x_{\pi(s)}, & t = 2s-2, \end{cases} \qquad A_t = \begin{cases} J_s, & t = 2s-1; \\ \{a_s\}, & t = 2s-2, \end{cases}$$

such that

$$J_1 = \langle a_1, a_2 \rangle, \dots, J_{l-1} = \langle a_{l-1}, a_l \rangle, J_l = \langle |a_l|, +\infty \rangle, J_{l+1} = \langle a_{l+2}, a_{l+1} \rangle, \dots, J_n = \langle a_{n+1}, a_n \rangle.$$

Proof. The regions A_t of the regression function are directly obtained by Corollary 2.3. So, the intervals J_s of the corresponding constant values $\alpha_t(\lambda) = x_{\pi(s)}$ are defined by statement (i) of Corollary 2.3. So, if $s \le l-1$ ($\delta_s = -1$), then

$$J_1 = \langle a_1, a_2 \rangle, \dots, J_{l-1} = \langle a_{l-1}, a_l \rangle,$$

while when $s \ge l+1$ ($\delta_s = 1$), it follows that

$$J_{l+1} = \langle a_{l+2}, a_{l+1} \rangle, \dots, J_n = \langle a_{n+1}, a_n \rangle.$$

Situation when $s = l$ the statement (iii) of Corollary 2.3 is satisfied, what directly implicates that

$$J_l = \langle |a_l|, +\infty \rangle.$$

Finally, taking into account statement (ii) of Corollary 2.3, the midpoints $\{a_s\}$ are defined, at which $\alpha_t(\lambda) = (1-\lambda)x_{\pi(s-1)} + \lambda x_{\pi(s)}$. □

Corollary 3.2. *It holds that*

$$\lim_{w \to +\infty} F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}) = x_k.$$

□

Corollary 3.3. *Let*

$$\lim_{w \to 0+} F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}) = x_{\pi(s_0)},$$

then it holds that:

(i) *if $s_0 \le l$, then $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$ monotonically increases;*
(ii) *if $s_0 \ge l+1$, then $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$ monotonically decreases.*

□

Corollary 3.4. *Let*

$$\lim_{w \to 0+} F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}) = x_{\pi(s_0)},$$

then it holds that:

(i) *if $s_0 \le l$ then*
   (a) *if $\lambda = 0$ then $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$ is left continuous;*
   (b) *if $\lambda = 1$ then $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$ is right continuous;*
(ii) *if $s_0 \ge l+1$ then*
   (c) *if $\lambda = 0$ then $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$ is right continuous;*
   (d) *if $\lambda = 1$ then $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$ is left continuous.*

□

Taking into account Proposition 2.4, the next corollary is presented directly.

Corollary 3.5. *Let $\alpha, \beta, \gamma \in \mathbb{R}$, such that $\alpha > 0$, $\beta \neq 0$, then it holds that*

$$F_k(\lambda, w; \alpha\boldsymbol{w}, \beta\boldsymbol{x} + \gamma\boldsymbol{e}) = \beta\, F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}) + \gamma,$$

where $\boldsymbol{e} = (1, \ldots, 1) \in \mathbb{R}^n$. □

Example 3.6. The next figure presents $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$ for $\boldsymbol{x} = (1, 2, 3, 4, 5)$ and some fixed value $\lambda \in [0, 1]$. Specifically, Figure 2(a) illustrates the case where $\boldsymbol{w} = (1, 2, 3, 2, w = w_5)$, $k = 5$, and $\lambda = 0$, while Figure 2(b) presents the situation for $\boldsymbol{w} = (w = w_1, 1, 3, 4, 1)$, $k = 1$, and $\lambda = 1$. The scenarios in the figures clearly illustrate properties that are directly derived from the definition of the weighted median. First, as $w \to +\infty$, it follows that $F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}) \to x_k$, a property presented in Corollary 3.2. Furthermore, the relationship between x_k and the element $x_{\pi(s_0)}$ is evident when $F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}) \to x_{\pi(s_0)}$ as $w \to 0+$, clearly demonstrating the effect on the monotonicity of the observed regression function, which is observed in Corollary 3.3. Furthermore, considering Corollary 3.4, it can be concluded that both functions satisfy the conditions of continuity on the left side.

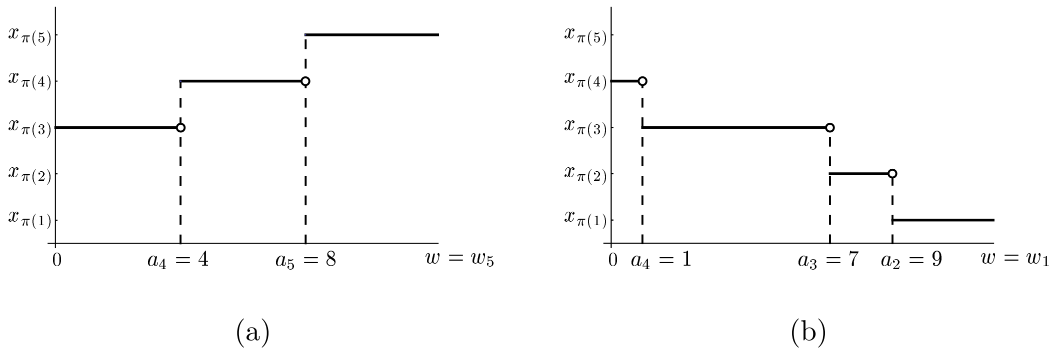


Figure 2. Graph of $w \mapsto F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x})$

3.2. The objective function analysis

The objective function Δ, as defined by (1.2), is generated by each corresponding regression function $F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j)$. According to Theorem 3.1, each $F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j)$ can be expressed as

$$F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j) = \sum_{t=1}^{2n-1} \alpha_t^{(j)}(\lambda)\, \mathbf{1}_{A_t^{(j)}}(w), \tag{3.1}$$

where $\alpha_t^{(j)}(\lambda)$ and $A_t^{(j)}$ are derived by the permutation $\pi_j \in S_n$, which sorts $\boldsymbol{x}_j \in \mathbb{R}^n$. Consequently, all possible regions $A_t^{(j)}$ must be obtained in order to describe the objective function Δ. So, let us define the set of midpoints that is generated by $\pi \in S_n$ as

$$\mathcal{A}_\pi = \{a_t(\pi; k, \boldsymbol{w}) \mid a_t(\pi; k, \boldsymbol{w}) > 0\}, \quad t \in \{1, \ldots, n\},$$

where $a_t = a_t(\pi; k, \boldsymbol{w})$ is defined by (2.4). By considering that Δ is generated by $m \gg 0$, the set of all possible midpoints can be written as

$$\mathcal{A}^* = \bigcup_{\pi \in S_n} \mathcal{A}_\pi, \tag{3.2}$$

and thus all possible regions of the objective function Δ are determined.

Proposition 3.7. *It holds that*

$$0 \leq |\mathcal{A}^*| \leq \sum_{t=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n-1}{t-1} - \frac{\operatorname{mod}(n,2)}{2} \cdot \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor}.$$

Proof. Considering (2.4), it can be concluded that $a_t > 0$ is obtained by substruction of two weight groups (where w_k is excluded). So, counting all $t-1$ combinations from $n-1$ elements, we can obtain the cardinal number $|\mathcal{A}^*|$.

First, consider the situation where n is an even number, i.e. $\operatorname{mod}(n,2) = 0$. In this situation the maximal number of positive coefficients $a_t > 0$ can be derived as

$$\begin{aligned} |\mathcal{A}^*| &\leq \binom{n-1}{0} + \binom{n-1}{1} + \cdots + \binom{n-1}{\frac{n}{2}-1} \\ &= \sum_{t=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n-1}{t-1}. \end{aligned} \tag{3.3}$$

Consider the situation where n is an odd number, that is, $\operatorname{mod}(n,2) = 1$. In this situation, it can be written that

$$\begin{aligned} |\mathcal{A}^*| &\leq \binom{n-1}{0} + \binom{n-1}{1} + \cdots + \frac{1}{2} \cdot \binom{n-1}{\frac{n-1}{2}} \\ &= \sum_{t=1}^{\lfloor \frac{n+1}{2} \rfloor} \binom{n-1}{t-1} - \frac{1}{2} \cdot \binom{n-1}{\lfloor \frac{n-1}{2} \rfloor}. \end{aligned} \tag{3.4}$$

The last part is divided because this situation occurs when two groups have the same number of elements, allowing them to generate double the coefficients. Finally, combining (3.3) and (3.4), we obtain the statement of the proposition. ☐

Proposition 3.8. *It holds that*

$$F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j) = \sum_{t=1}^{2r+1} \beta_t^{(j)}(\lambda)\, \mathbf{1}_{A_t^*}(w),$$

where

$$A_t^* = \begin{cases} J_s^*, & t = 2s-1; \\ \{a_s^*\}, & t = 2s, \end{cases} \qquad t \in \{1, \ldots, 2r+1\},$$

such that J_s^* *are intervals defined as*

$$J_1^* = \langle 0, a_1^* \rangle, \ldots, J_s^* = \langle a_{s-1}^*, a_s^* \rangle, \ldots, J_{r+1}^* = \langle a_r^*, +\infty \rangle.$$

where

$$0 < a_1^* < a_2^* < \cdots < a_r^* < +\infty, \quad r = |\mathcal{A}^*|,$$

denotes the sorted elements of $\mathcal{A}^*$*, and it holds that* $\beta_t^{(j)}(\lambda) = \alpha_s^{(j)}(\lambda)$ *iff* $A_t^* \subseteq A_s^{(j)}$*.*

Proof. According to (3.2) it holds that $\mathcal{A}_{\pi_j} \subseteq \mathcal{A}^*$, so for every $A_s^{(j)}$ of (3.1) there exists A_t^* such that $A_t^* \subseteq A_s^{(j)}$, and thus $\beta_t^{(j)}(\lambda) = \alpha_s^{(j)}(\lambda)$. ☐

Theorem 3.9. *It holds that*

$$\Delta(\lambda, w) = \sum_{t=1}^{2r+1} \alpha_t^*(\lambda)\ \mathbf{1}_{A_t^*}(w),$$

where

$$\alpha_t^*(\lambda) = \sum_{j=1}^{m} |y_j - \beta_t^{(j)}(\lambda)|^p.$$

Proof. According to Proposition 3.8, each regression function $F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j)$ defined by (3.1) can be written as the piecewise function that is defined in regions A_t^*. So, it can be written that

$$\begin{aligned}\Delta(\lambda, w) &= \sum_{j=1}^{m} |y_j - F_k(\lambda, w; \boldsymbol{w}, \boldsymbol{x}_j)|^p,\\ &= \sum_{j=1}^{m} |y_j - \sum_{t=1}^{2r+1} \beta_t^{(j)}(\lambda)\, \mathbf{1}_{A_t^*}(w)|^p,\\ &= \sum_{t=1}^{2r+1}\sum_{j=1}^{m} |y_j - \beta_t^{(j)}(\lambda)|^p\, \mathbf{1}_{A_t^*}(w),\end{aligned}$$

what proves the statement of the theorem. □

4. Optimization of the component weighted median model

Referring to Section 3, the properties of the observed model (1.1) have been studied and presented, allowing optimization to be performed.

Theorem 4.1. *It holds that*

$$\min_{\substack{\lambda\in[0,1]\\ w>0}} \Delta(\lambda, w) = \min_{1\le t\le 2r+1} \Delta(\lambda_t^*, w_t^*),$$

where $w_t^* \in A_t^*$ *and*

$$\lambda_t^* = \operatorname*{argmin}_{\lambda\in[0,1]} \Delta(\lambda, w_t^*).$$

Proof. According to Theorem 3.9, the objective function Δ is constant for $w_t^* \in A_t^* = J_s^*$ and any $\lambda_t^* \in [0,1]$, while at the midpoints, i.e., when $w_t^* \in A_t^* = \{a_s^*\}$, $t = 2s$, the corresponding parameter λ_t^* can be optimized. In this situation, due to $\lambda \mapsto \Delta(\lambda, a_s^*)$ being continuous and convex, the global minimum λ_t^* can always be reached. So, it can be concluded that the optimization problem of Δ can be presented discretely. □

Remark 4.2. By considering Theorem 2.1, it may be concluded that the parameter λ_t^* for the corresponding midpoint region $A_t^* = \{a_s^*\}$ can be explicitly expressed as

$$\lambda_t^* = \begin{cases} \operatorname{med}_\lambda(\widetilde{\boldsymbol{w}}_t, \widetilde{\boldsymbol{x}}_t), & 0 \le \operatorname{med}_\lambda(\widetilde{\boldsymbol{w}}_t, \widetilde{\boldsymbol{x}}_t) \le 1;\\ 0, & \operatorname{med}_\lambda(\widetilde{\boldsymbol{w}}_t, \widetilde{\boldsymbol{x}}_t) < 0;\\ 1, & \operatorname{med}_\lambda(\widetilde{\boldsymbol{w}}_t, \widetilde{\boldsymbol{x}}_t) > 1,\end{cases}$$

such that

$$\widetilde{\boldsymbol{w}}_t = \Big(F_k(1, a_s^*; \boldsymbol{w}, \boldsymbol{x}_{g(j)}) - F_k(0, a_s^*; \boldsymbol{w}, \boldsymbol{x}_{g(j)}) \mid j = 1, \ldots, m_t\Big),$$

$$\widetilde{\boldsymbol{x}}_t = \left(\frac{y_{g(j)} - F_k(0, a_s^*; \boldsymbol{w}, \boldsymbol{x}_{g(j)})}{F_k(1, a_s^*; \boldsymbol{w}, \boldsymbol{x}_{g(j)}) - F_k(0, a_s^*; \boldsymbol{w}, \boldsymbol{x}_{g(j)})} \mid j = 1, \ldots, m_t\right),$$

where $g\colon \{1, \ldots, m_t\} \to \mathcal{T}_t$, $m_t = |\mathcal{T}_t|$, denotes some mapping on the sets of indices

$$\mathcal{T}_t = \{j \mid F_k(0, a_s^*; \boldsymbol{w}, \boldsymbol{x}_j) \ne F_k(1, a_s^*; \boldsymbol{w}, \boldsymbol{x}_j)\}, \quad j \in \{1, \ldots, m\}.$$

Remark 4.3. The computational efficiency of the proposed optimization model is based on the average computational complexity of the Quicksort algorithm. In this context, it can be concluded that the evaluation of the objective function Δ requires the classification of m vectors of size n, which involves $\mathcal{O}(mn\log n)$ operations. Furthermore, evaluating Δ for each A_t^* requires $\mathcal{O}(m(2r+1))$ operations, where the complexity of generating the observed regions is discussed in Proposition 3.7. Furthermore, according to Remark 4.2, determining the parameters λ_t^* in the midpoint region involves a computational complexity of $\mathcal{O}(r\widetilde{m}\log\widetilde{m})$ operations, where $\widetilde{m} \le m$.

5. Numerical example and application

In this section, we present the performance and application of the observed model, Δ. We specifically address the problem of expectation estimation for normal distributions, using it as a numerical example while considering outliers in the low-probability tails. Our application focus is on the image denoising problem, a critical area in image processing [3, 16]. This topic is especially important for the reduction of noise in corrupted medical images, as effective denoising is essential for an accurate medical diagnosis. The methods discussed emphasize median-based fuzzy approaches [6, 21]. Therefore, we investigate a generalized model to optimize well-known adaptive impulse removal techniques that aggregate the sums of objective functions across different input dimensions and weighting positions.

5.1. Numerical example

To demonstrate the optimization of the observed model Δ, we consider a normally distributed data vector $\boldsymbol{x}_j = (x_1^{(j)}, \ldots, x_n^{(j)})$, where $x_i^{(j)} \sim \mathcal{N}(\mu_j, \sigma_j^2)$ and $y_j = \mu_j$. This scenario also includes outliers in $\boldsymbol{x}_j$, which are typically found in the low-probability tails. Such outliers can lead to model misspecification and yield inaccurate results. Therefore, the robustness of the median in relation to outliers helps mitigate these issues by preserving the integrity of the model [17]. Figure 3 presents the optimization model that is performed considering Theorem 4.1. Performance is demonstrated for $n = 9$ and $m = 10$ where the center weight $w = w_5$ is observed, while $w_i = 1$, $i \neq 5$. The red-marked graph presents a discreet optimization path across the regions of the objective function Δ, where the midpoint region is optimized considering Remark 4.2.

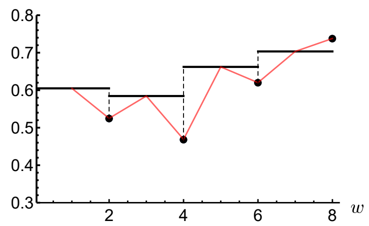


Figure 3. Optimization model

Figure 4 presents the cases of a normal distribution that are generated in the observed model case, which reach its global minimum at $(\lambda^*, w^*) = (1, 4)$ (Figure 3). In these cases, outliers are indicated as red points and regular data as black, while the green point denotes weighted data $x_5^{(j)}$. Figures 4(a) and (b) demonstrate the advantage of the observed model, where the green dashed line presents the position of $F_k(\lambda^*, w^*; \boldsymbol{w}, \boldsymbol{x}_j)$, $k = 5$, which effectively estimate expectation $y_j = \mu_j$. In other situations, Figures 4(c) and (d) demonstrate the disadvantages of the weight positioning and outliers, which lead to a poor estimation of the corresponding expectation.

5.2. Image denoising

This subsection discusses the application of model (1.1), which is commonly used in the field of image processing, particularly to address the image denoising problem [3, 16]. In this field, various types of filters are developed to target specific types of noise that degrade the visual quality of images. The image of size $N \times M$ in this area is represented matrix-wise as $\mathtt{Y} = [\mathtt{y}_{\mathtt{i},\mathtt{j}}]_{1 \le \mathtt{i} \le N, 1 \le \mathtt{j} \le M}$, where each component $\mathtt{y}_{\mathtt{i},\mathtt{j}}$ represents the color intensity of the corresponding pixel positions. One of the possible representations of the pixel color intensities considered in this work is the representation in grayscale levels. In this scenario,

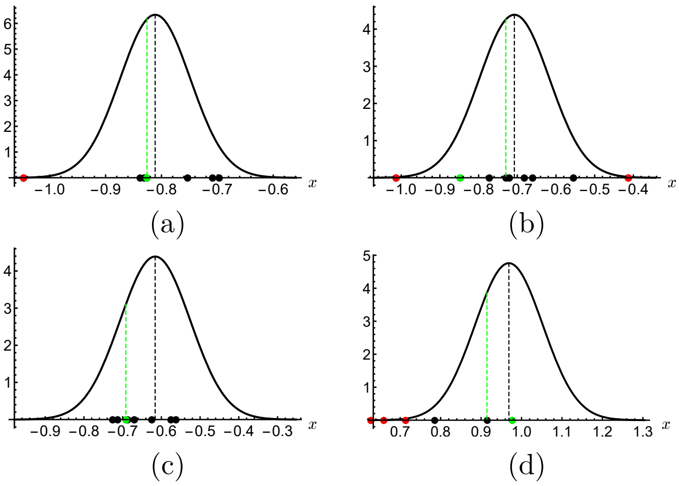


Figure 4. Optimization results

the scale of the color intensities is defined such that $\mathtt{y}_{\mathtt{i},\mathtt{j}} \in [0,1]$, where 0 indicates black and 1 indicates white, while the intermediate values represent different shades of gray. In this field, various mathematical models are used to simulate different types of noise present in images. In this investigation, the model that describes impulse noise, also known as 'salt and pepper' noise, which generates a noisy image $\mathtt{X} = [\mathtt{x}_{\mathtt{i},\mathtt{j}}]_{1\leq \mathtt{i},\leq N, 1\leq \mathtt{j}\leq M}$, such that

$$\mathtt{x}_{\mathtt{i},\mathtt{j}} = \begin{cases} \xi_{\mathtt{i},\mathtt{j}}, & \text{with probability } \rho; \\ \mathtt{y}_{\mathtt{i},\mathtt{j}}, & \text{with probabilitiy } 1-\rho. \end{cases} \tag{5.1}$$

In the observed model $\rho \in [0,1]$ represents the noise ratio, while $\xi_{\mathtt{i},\mathtt{j}}$ represents a random variable whose probability density function is defined as

$$P(\mathtt{t}) = \begin{cases} P_{\mathtt{p}}, & \mathtt{t} = \mathtt{p}; \\ P_{\mathtt{s}}, & \mathtt{t} = \mathtt{s}; \\ 0, & \text{otherwise}, \end{cases}$$

where $P_{\mathtt{p}}, P_{\mathtt{s}} \geq 0$ $(P_{\mathtt{p}} + P_{\mathtt{s}} = 1)$ represent the probabilities of the occurrence of values $\mathtt{p}$ and $\mathtt{s}$, respectively. For 'salt and pepper' noise, it is common to use the minimum and maximum values of the color intensity scale with the same probability of occurrence. Therefore, the corresponding values $\mathtt{p} = 0$ and $\mathtt{s} = 1$ are considered, while the probabilities of occurrence are equal, that is, $P_{\mathtt{p}} = P_{\mathtt{s}} = 0.5$. It is well-known that median-based filters are primarily used to suppress impulse noise corruption in images. This effectiveness comes from its robustness to outliers, which in this case represent impulse noise [9, 16, 17]. Most of these methods are constructed to operate on the filtering window of each $\mathtt{x}_{\mathtt{i},\mathtt{j}}$, resulting in a new reconstructed image $\mathtt{X}^* = [\mathtt{x}^*_{\mathtt{i},\mathtt{j}}]_{1\leq \mathtt{i}\leq N, 1\leq \mathtt{j}\leq M}$. A standard 3×3 filtering window, which is a commonly used size for filtering, is typically represented in vector form, and is presented as follows.

$$\begin{array}{|c|c|c|} \hline \mathtt{x}_{\mathtt{i}-1,\mathtt{j}-1} & \mathtt{x}_{\mathtt{i},\mathtt{j}-1} & \mathtt{x}_{\mathtt{i}+1,\mathtt{j}-1} \\ \hline \mathtt{x}_{\mathtt{i}-1,\mathtt{j}} & \mathtt{x}_{\mathtt{i},\mathtt{j}} & \mathtt{x}_{\mathtt{i}+1,\mathtt{j}} \\ \hline \mathtt{x}_{\mathtt{i}-1,\mathtt{j}+1} & \mathtt{x}_{\mathtt{i},\mathtt{j}+1} & \mathtt{x}_{\mathtt{i}+1,\mathtt{j}+1} \\ \hline \end{array} = \begin{array}{|c|c|c|} \hline x_1^{(j)} & x_2^{(j)} & x_3^{(j)} \\ \hline x_4^{(j)} & x_5^{(j)} & x_6^{(j)} \\ \hline x_7^{(j)} & x_8^{(j)} & x_9^{(j)} \\ \hline \end{array}$$

In this situation, it can be denoted that $\boldsymbol{x}_j = (x_1^{(j)}, \ldots, x_9^{(j)})$, while $y_j = \mathtt{y}_{\mathtt{i},\mathtt{j}}$, where $j = (\mathtt{i}-1)M + \mathtt{j}$.

5.2.1. Improved denoising of MR data in image processing. Digital impulses mixed with Magnetic Resonance (MR) images can render them unusable for diagnosis, leading to unrecognizable images that contain misleading information. This noise may occur during the reconstruction phase or originate from external radio frequency waves. Therefore, it is essential to remove this noise from MR images [21]. To address this issue, an investigation is conducted on different types of methods that effectively remove impulse noise from images, where the observed model is implemented to improve filtering performance. Furthermore, due to the complexity of the observed model, the research focuses only on pixels identified as noisy, while others are treated as non-noisy based on a specific detection procedure of the filtering method and thus left unchanged. Among the methods studied, SAMF [5] (Simple Adaptive Median Filter), AMF [8, 21] (Adaptive Median Filter), and NAFSM [20] (Noise Adaptive Fuzzy Switching Median Filter) incorporate flexible window sizes. In contrast, other methods such as CWMF [7] (Central Weighted Median Filter), SCWMF [19] (Switching Central Weighted Median Filter), and SMF [23] (Switching Median Filter) do not include these types of flexibility, with CWMF being particularly noted for lacking any noise detectors and thus processing all pixels.

The performance of the filtering methods is presented using the Absolute Deviations Error (ADE) measure, defined in this context as

$$\mathrm{ADE} = \sum_{\mathtt{i}=1}^{M}\sum_{\mathtt{j}=1}^{N} |\mathtt{y}_{\mathtt{i},\mathtt{j}} - ((1-\mathtt{n}_{\mathtt{i},\mathtt{j}})\mathtt{x}_{\mathtt{i},\mathtt{j}} + \mathtt{n}_{\mathtt{i},\mathtt{j}}\mathtt{x}^*_{\mathtt{i},\mathtt{j}})|,$$

where the general noise detector scheme is represented as

$$\mathtt{n}_{\mathtt{i},\mathtt{j}} = \begin{cases} 1, & \mathtt{x}_{\mathtt{i},\mathtt{j}} \text{ is noisy;} \\ 0, & \mathtt{x}_{\mathtt{i},\mathtt{j}} \text{ is no-noisy.} \end{cases}$$

In this way, unnecessary filtering is controlled, managing the replacement of non-noisy pixels, which preserves fine details and prevents edge distortion. To achieve this, the observed component of the weighted median model is generalized to incorporate all specific filtering modifications to improve filtering performance. Thus, a certain modification is carried out through the generalized objective function:

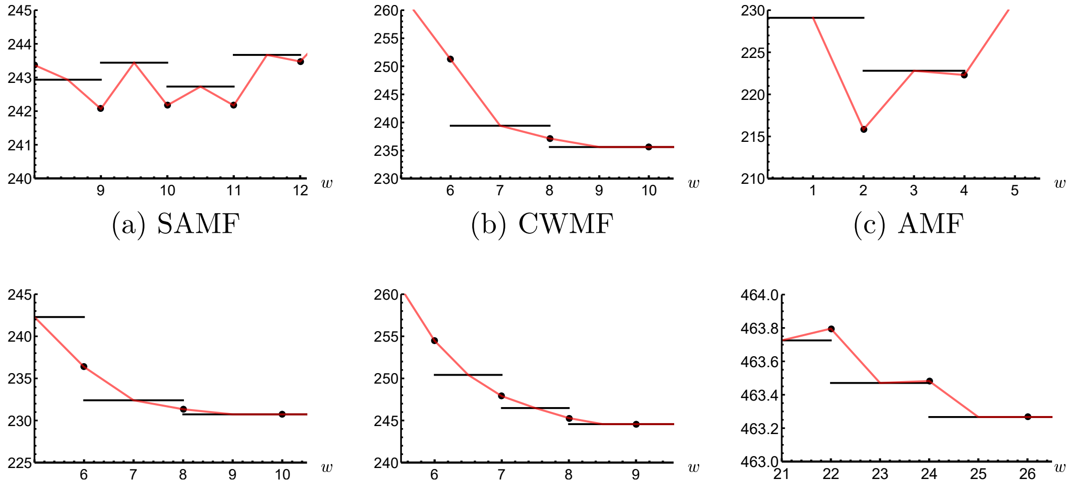
$$\widetilde{\Delta}(\lambda, w) = \sum_{j=1}^{\widetilde{m}} |\widetilde{y}_j - F_{k_j}(\lambda, w; \widetilde{\boldsymbol{w}}_j, \widetilde{\boldsymbol{x}}_j)|,$$

which relates only to those observations considered noisy by the specified noise detectors. Specifically, $\widetilde{y}_j = \mathtt{y}_{\mathtt{i},\mathtt{j}}$, while $\widetilde{\boldsymbol{x}}_j \in \mathbb{R}^{n_j}$ corresponds to the filtering neighborhood of $\mathtt{x}_{\mathtt{i},\mathtt{j}}$, which is detected as noisy, i.e., when $\mathtt{n}_{\mathtt{i},\mathtt{j}} = 1$. Furthermore, the specific weighting position k_j is observed ($w^{(j)}_{k_j} = w$), where the first non-noisy pixel closest to the filtering window center is considered, and all other weights are set to one, i.e., $w^{(j)}_i = 1$ for $i \neq k_j$. This process is conducted by using the simple impulse noise detector:

$$\widetilde{\mathtt{n}}_{\mathtt{i},\mathtt{j}} = \begin{cases} 1, & \text{if } \mathtt{x}_{\mathtt{i},\mathtt{j}} = 0 \text{ or } \mathtt{x}_{\mathtt{i},\mathtt{j}} = 1; \\ 0, & \text{else.} \end{cases}$$

This approach effectively detects non-noisy pixels within the observed filtering window $\widetilde{\boldsymbol{x}}_j \in \mathbb{R}^{n_j}$, while all other noisy pixels are excluded from the filtering process [5, 20]. Consequently, this method avoids replacing noisy pixels, thereby preserving fine details and edges in the image and preventing blurring effects. Consequently, by Theorem 3.9, the regions of $\widetilde{\Delta}$ can be derived by observing the sums of objective functions with the same input dimensions n and position k (i.e., $\Delta_{k,n}$). In this manner, it follows that $\widetilde{\Delta} = \sum_{k,n} \Delta_{k,n}$, where the corresponding midpoint set of $\Delta_{k,n}$ is denoted as $\mathcal{A}^*_{k,n}$, resulting in the final midpoint set of $\widetilde{\Delta}$ given by $\widetilde{\mathcal{A}}^* = \cup_{k,n}\mathcal{A}^*_{k,n}$.

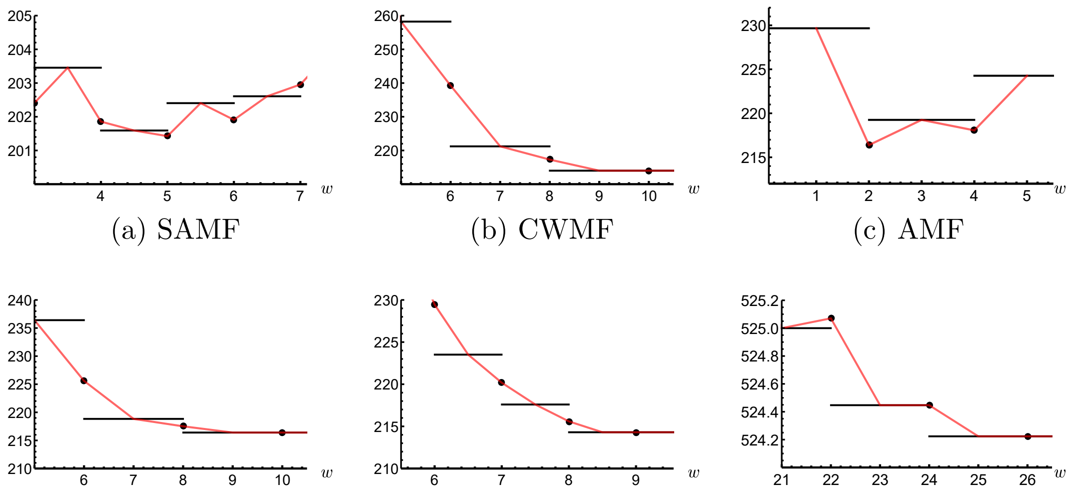
Therefore, the generalized objective model $\widetilde{\Delta}$ is examined to improve the proposed filtering methods. Thus, Figures 5 and 6 present the filtering results through ADE for each corresponding filtering model modified by $\widetilde{\Delta}$, conducted on two distinct test images: MR1 and MR2.



(a) SAMF (b) CWMF (c) AMF

(d) SCWMF (e) NAFSM (f) SMF

Figure 5. Filtering results for MR1



(a) SAMF (b) CWMF (c) AMF

(d) SCWMF (e) NAFSM (f) SMF

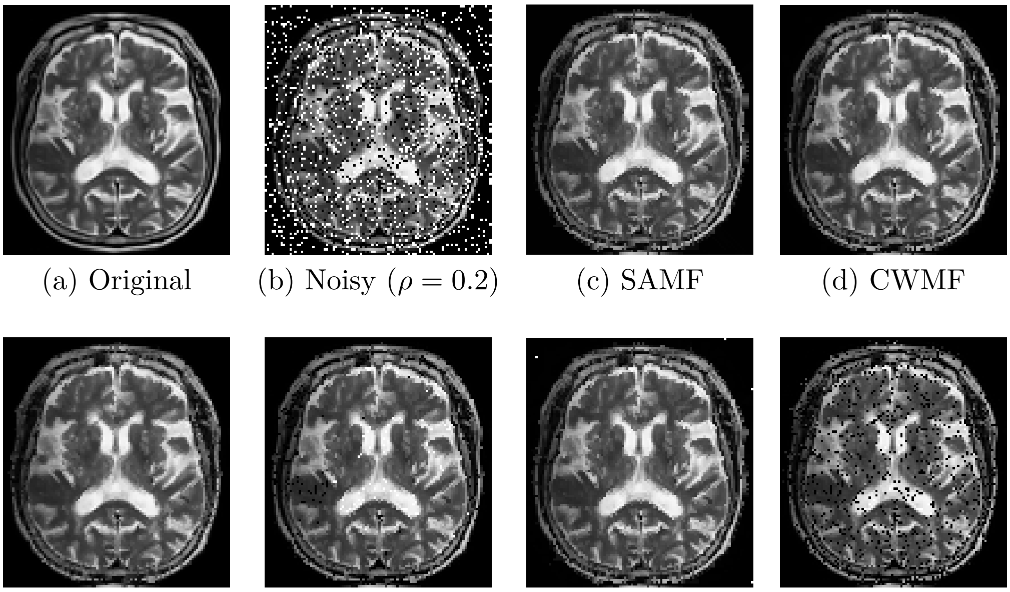
Figure 6. Filtering results for MR2

In Table 1, the performance of the standard and optimized methods is compared. It is shown that the optimized SAMF significantly outperforms the standard variation, as the filtering window increases until a certain number of non-noisy detected pixels is reached, at which point the standard median is used for the filtering output [5]. The optimized parameters for SAMF corresponding to $\widetilde{\Delta}$ are achieved at the midpoint $(\lambda^*, w^*) = (0, 9)$ for the MR1 image (Figure 5(a)), and at $(\lambda^*, w^*) = (0.454, 5)$ for MR2 (Figure 6(a)). The results for the CWMF indicate that the proposed modification significantly surpasses the standard variation. The standard CWMF processes every pixel using a 3×3 filtering window, where the center weight is $w_5 = 3$ and $w_i = 1$ for $i \neq 5$ [7], while the proposed modification introduces flexible weighting. The optimal performance of the modified CWMF is attained at $(\lambda^*, w^*) \in [0,1] \times \langle 8, \infty\rangle$ for both MR test images (Figures 5(b) and 6(b)). The AMF increases the filtering window until the median is not found between the extreme values within the current window. Subsequently, the output of the standard AMF is defined as the median of the observed window [8, 21]. The proposed modification demonstrates significant improvements, with optimized parameters of $(\lambda^*, w^*) = (0.382, 2)$ for MR1 (Figure 5(c)) and $(\lambda^*, w^*) = (0.390, 2)$ for MR2 (Figure 6(c)). The standard variation of SCWMF employs a 3×3 filtering window for noise detection, incorporating the distance of the observed pixel to the median value of the window. The output is determined as the center-weighted median, where $w_5 = 3$ and $w_i = 1$ for $i \neq 5$ [19]. The proposed modification shows significant improvements, with optimized parameters for both MR test images achieved at $(\lambda^*, w^*) \in [0,1] \times \langle 8, \infty\rangle$ (Figures 5(d) and 6(d)). The standard variation of NAFSM incorporates a fuzzy noise detector, with the filtering window expanding until one non-noisy pixel is detected. Subsequently, the standard median of the noisy detected pixels within the specified filtering window is used as the output [20]. The results show that the proposed model significantly improves the standard NAFSM, with optimized parameters for both MR test images reaching $(\lambda^*, w^*) \in [0,1] \times \langle 8, \infty\rangle$ (Figures 5(e) and 6(e)). The standard variation of SMF involves detecting noisy pixels by considering the minimum of four types of convolution applied to a 5×5 filtering window, where the output is defined as the standard median of the observed filtering window [23]. It is evident that the proposed model significantly improves the standard method, with optimized parameters for both MR test images reaching $(\lambda^*, w^*) \in [0,1] \times \langle 24, \infty\rangle$ (Figures 5(f) and 6(f)).

Table 1. ADE measures

Method	Standard	Optimized	Outperformance
	MR1 ($\rho = 20\%$)		
SAMF [5]	322.282	242.067	24.89%
CWMF [7]	410.933	235.616	42.66%
AMF [8, 21]	229.094	215.837	5.79%
SCWMF [19]	339.400	230.729	32.02%
NAFSM [20]	529,992	244.568	53.85%
SMF [23]	515.510	463.267	10.13%
	MR2 ($\rho = 20\%$)		
SAMF [5]	213.259	201.425	5.55%
CWMF [7]	429.024	214.004	50.12%
AMF [8, 21]	229.663	216.384	5.78%
SCWMF [19]	339.055	216.400	36.18%
NAFSM [20]	544.331	214.301	60.63%
SMF [23]	601.784	524.224	12.89%

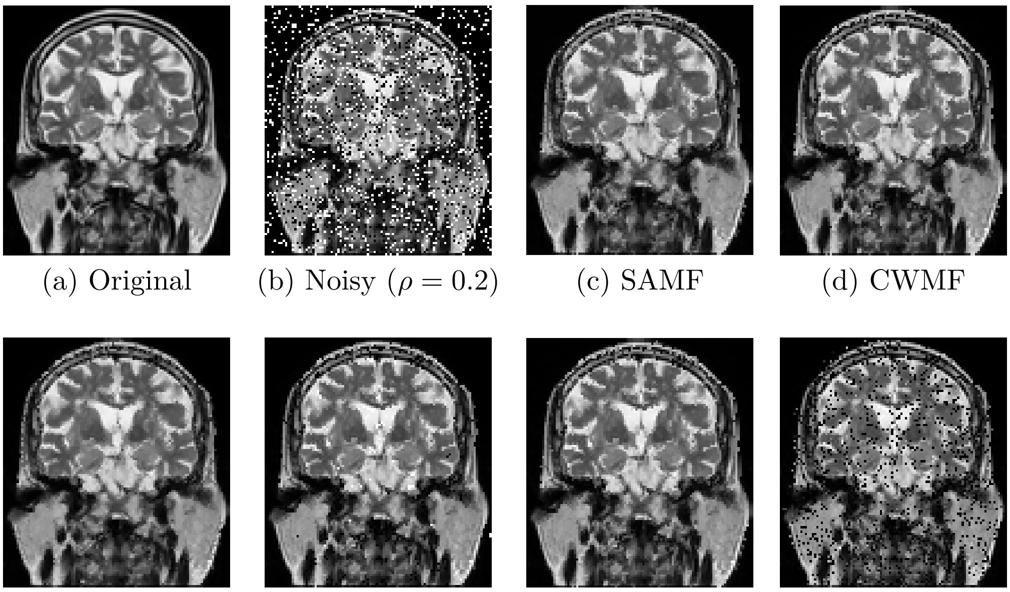
Figures 7 and 8 present the optimal filtering results for the MR1 and MR2 images, both of which have the same size of 110×100. Figures 7(a) and 8(a) present the original MR1 and MR2 images, respectively, where the original MR1 shows the axial view that captures a horizontal image from the top of the body to the bottom, while the original MR2 displays the coronal viewa frontal view that presents a mirror image of the body from front to back. Figures 7(b) and 8(b) present the noisy MR1 and MR2 images, which are generated by the impulse noise model (5.1) with $\rho = 0.2$. In this context, the noisy images MR1 and MR2 demonstrate the characteristics of impulse noise, which is defined by random occurrences of bright (salt) and dark (pepper) pixels, resembling salt and pepper sprinkled on an image. As illustrated, all the presented filters effectively reconstruct the noisy images, with the AMF demonstrating the best performance for the MR1 image, while the SAMF excels for the MR2 image.



(a) Original (b) Noisy ($\rho = 0.2$) (c) SAMF (d) CWMF

(e) AMF (f) SCWMF (g) NAFSM (h) SMF

Figure 7. Optimal filtering results for MR1



(a) Original (b) Noisy ($\rho = 0.2$) (c) SAMF (d) CWMF

(e) AMF (f) SCWMF (g) NAFSM (h) SMF

Figure 8. Optimal filtering results for MR2

6. Conclusion

This study presents a comprehensive analysis of a multidimensional regression model based on the weighted median, demonstrating that the associated optimization problem can be effectively discretized. The detailed examination of the regression and objective functions reveals their piecewise nature, providing a foundation for the efficient computation of the optimal solution. The robustness of the proposed L_1 norm model to outliers, particularly its effectiveness in image denoising applications, has been demonstrated, showing that incorporating the model into adaptive filtering techniques yields notable improvements in removing impulse noise from MR images. The results suggest that the presented methodology provides a valuable framework for robust regression analysis and image denoising. Future research could explore the application of this model to even more sophisticated denoising techniques and also apply it to other areas where the presence of outliers can severely affect the accuracy of the model.

Acknowledgements

The authors sincerely thank the editorial and reviewer team of the Hacettepe Journal of Mathematics and Statistics for their helpful and clear suggestions. Their insightful comments and constructive feedback have significantly improved our paper in all aspects.

Author contributions. All the co-authors have contributed equally in all aspects of the preparation of this submission.

Conflict of interest statement. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding. This work is not supported by any funding.

Data availability. The data utilized in the research presented in this article are accessible at www.pockethealth.com.

References

[1] P. Bloomfield and W. L. Steiger, *Least Absolute Deviations: Theory, Applications, and Algorithms*, Birkhäuser, Boston, 1983.

[2] Y. Dodge, *An introduction to L_1-norm based statistical data analysis*, Comput. Stat. Data Anal.**5**(4), 239253, 1987.

[3] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Third Edition, Pearson Prentice Hall, New Jersey, 2008.

[4] C. Gurwitz, *Weighted median algorithms for L_1 approximation*, BIT **30**(2), 301-310, 1990.

[5] H. Ibrahim, N. S. Pik Kong and T. F. Ng, *Simple adaptive median filter for the removal of impulse noise from highly corrupted images*, IEEE Trans. Consum. Electron. **54**(4), 1920-1927, 2008.

[6] C. J. J. Sheela and G. Suganthi, *An efficient denoising of impulse noise from MRI using adaptive switching modified decision based unsymmetric trimmed median filter*, Biomed. Signal Process. Control **55**, 101657, 2020.

[7] S. J. Ko and Y. H. Lee, *Center Weighted Median filters and their applications and image enhancements*, IEEE Trans. Circuits syst. **38**(9), 984-993, 1991.

[8] G. Y. Lee, G. L. Ra, G. L. Ra, G. S. Kim, H. H. Moon and J. S. Jeong, *Probability Mass Function-Based Adaptive Median Filtering for Acoustic Radiation Force Impulse Imaging: A Feasibility Study*, IEEE Access **11**(11), 2169-3536, 2023.

[9] M. M. Ali, R. Imon, I. Ali and H. M. Yousof, *Statistical Outliers and Related Topics*, CRR Press, Taylor & Francis Group, 2025.

[10] V. Novoselac, *The Component Weighted Median Absolute Deviations Problem*, Eur. J. Pure Appl. Math. **13**(4), 964-976, 2020.

[11] V. Novoselac and Z. Pavić, *Cluster detection in noisy environment by using the modified EM algorithm*, Croat. Oper. Res. Rev. **9**(2), 223-234, 2018.

[12] V. Novoselac and Z. Pavić, *Optimal Solution Properties of an Overdetermined System of Linear Equations*, Eur. J. Pure Appl. Math. **12**(4), 1360-1370, 2019.

[13] M. R. Osborne, *Finite Algorithms in Optimization and Data Analysis*, Department of Statistics, Australian National University, Camberra, John Wiley, 1985.

[14] Y. W. Park, *Optimization for L_1-Norm Error Fitting via Data Aggregation*, INFORMS J. Comput. **33**(1), 120-140, 2021.

[15] Z. Pavić and V. Novoselac, *Investigating an overdetermined system of linear equations by using convex functions*, Hacet. J. Math. Stat. **46**(5), 865-874, 2017.

[16] P. Qiu, *Image Processing and Jump Regression Analysis*, John Wiley & Sons, 2005.

[17] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, Wiley, New York, 2003.

[18] K. Sabo and R. Scitovski, *The best least absolute deviations line-properties and two efficient methods for its derivation*, Anziam J. **50**(2), 185-198, 2008.

[19] T. Sun and Y. Neuvo, *Detail-preserving median based filters in image processing*, Pattern Recognit. Lett. **15**(4), 341-347, 1994.

[20] K. K. V. Toh and N. A. Mat Isa, *Noise Adaptive Fuzzy Switching Median Filter for Salt-and-Pepper Noise Reduction*, IEEE Signal Process. Lett. **17**(3), 281-284, 2010.

[21] A. Toprak and . Güler, *Suppression of Impulse Noise in Medical Images with the Use of Fuzzy Adaptive Median Filter.* J. Med. Syst., **30**(6), 465471, 2006.

[22] I. Vazler, K. Sabo and R. Scitovski, *Weighted median of the data in solving least absolute deviations problems*, Comm. Statist. Theory Methods **41**(8), 1455-1465, 2012.

[23] S. Zhang and M. A. Karim, *A new impulse detector for switching median filters*, IEEE Signal Process. Lett. **9**(11), 360-363, 2002.