ECONOMETRICS
EKONOMETRIA
Advances in Applied Data Analysis

Econometrics. Ekonometria.
Advances in Applied Data Analysis
Year 2024, Vol. 28, No. 3
ISSN 2449-9994
journals.ue.wroc.pl/eada

A United Probabilistic Approach to Minimising the Sum of Absolute Values

Changyong Feng

University of Rochester, USA

E-mail: changyong_feng@urmc.rochester.edu

ORCID: 0000-0002-4432-1565

Honghong Liu

University of Rochester, USA

E-mail: Honghong_liu@urmc.rochester.edu

ORCID: 0009-0006-9260-2384

Ethan Poon

Cistercian Preparatory School, USA

E-mail: ethanchpoon@gmail.com

Ge Feng

Jianghan University, China

E-mail: fengge068@foxmail.com

Quote as: Feng, C., Liu, H., Poon, E., & Feng, G. (2024). A United Probabilistic Approach to Minimising the Sum of Absolute Values*. Econometrics. Ekonometria. Advances in Applied Data Analysis*, *28*(3), 1-8.

DOI: 10.15611/eada.2024.3.01

JEL: C01, C10, C18

Abstract

Aim: Introduce a novel method for minimizing functions in the form of a sum of absolute values.

Methodology: The sum of absolute values can be standardized so that the sum of the coefficients equals 1. In this case, the sum of absolute values takes the form $E|X - a|$, where X is a random variable.

Results: Any median of X is a minimiser of the function $E|X-a|$. To minimise the function, it suffices to find any median of X.

Implications and recommendations: The method introduced in this paper can be applied to minimise a large family of functions.

Originality/value: Our work uses the probabilistic method to solve optimization problems.

Keywords: median, probability distribution, minimum value

1. Introduction

To find the minimum of the following function (Xu, 2012, p. 54, Example 1):

$$f(x) = |x| + 2|x-1| + |x-2| + |x-4| + |x-6| + 2|x-10|,\ x \in \mathbb{R}. \tag{1}$$

Before approaching this problem, let us study an easier one. Suppose *a* and *b* are two constants and *a* < *b*. Consider function:

$$f(x) = |x-a| + |x-b|,\ x \in \mathbb{R}. \tag{2}$$

Lemma 1. *An x minimises f in (2) if and only if $x \in [a,b]$.*

Proof. One can write f as

$$f(x) = \begin{cases} (a+b) - 2x & \text{if } x \le a, \\ b-a & \text{if } a < x \le b, \\ 2x - (a+b) & \text{if } b < x. \end{cases}$$

The minimum of f is $b - a$, and $f(x)$ takes this value if and only if $x \in [a,b]$.

Lemma 1 indicates that the set of minimisers of f in (2) is exactly the closed interval $[a,b]$. In particular, f achieves its minimum at both endpoints of the interval.

Theorem 2. *Given the sequence of real numbers $a_1 < a_2 < \cdots < a_n$, let*

$$f(x) = |x-a_1| + |x-a_2| + \cdots + |x-a_n|,\ x \in \mathbb{R}. \tag{3}$$

(a) If n is odd, f achieves its minimum at $x = a_{(n+1)/2}$.

(b) If n is even, f achieves its minimum at any $x \in [a_{n/2}, a_{(n+2)/2}]$.

The following proof is of Nahin (2004). One can write f as

$$f(x) = (|x-a_1| + |x-a_n|) + (|x-a_2| + |x-a_{n-1}|) + \dots$$

From Lemma 1, if x minimises the sum within each pair of parentheses, it will minimise f in (3). If n is odd, only $a_{(n+1)/2}$ does the work. If n is even, any $x \in [a_{n/2}, a_{(n+2)/2}]$ works. Note that $a_{(n+1)/2}$ [respectively, $(a_{n/2} + a_{(n+2)/2})/2$] is the median of the sequence $a_1, a_2, \dots, a_n$ when n is odd (respectively, even). In the following sections we show that if n is event, it is more meaningful to designate any point within $[a_{n/2}, a_{(n+2)/2}]$ as a median of the sequence.

Suppose f is of the form:

$$f(x) = m_1|x-a_1| + m_2|x-a_2| + \cdots + m_n|x-a_n|,\ x \in \mathbb{R},$$

where $m_i, i = 1, \dots, n$, are positive integers, and $a_1 < a_2 < \cdots < a_n$. Using the same method as above one can prove that f achieves its minimum at any median of the following sequence:

$$\underbrace{a_1, \dots, a_1}_{m_1}, \quad \underbrace{a_2, \dots, a_2}_{m_2}, \dots, \quad \underbrace{a_n, \dots, a_n}_{m_n}.$$

The function at the beginning of this section can be written as:

$$f(x) = |x| + |x-1| + |x-1| + |x-2| + |x-4| + |x-6| + |x-10| + |x-10|.$$

Considering the sequence:

$$0,\ 1,\ 1,\ 2,\ 4,\ 6,\ 10,\ 10,$$

one can see that f achieves its minimum at any $x \in [2,4]$.

A generalisation of Theorem 2, where coefficients are positive real numbers, has also been discussed in the literature (cf. Bogomolny (2021)).

Theorem 3. *Given a sequence of real numbers $a_1 < a_2 < \cdots < a_n$, and a sequence of positive numbers $\beta_1, \beta_2, \dots, \beta_n$, let*

$$f(x) = \beta_1|x-a_1| + \beta_2|x-a_2| + \cdots + \beta_n|x-a_n|, x \in \mathbb{R}. \tag{4}$$

Then f achieves its minimum at one of the intervals $[a_k, a_{k+1}]$ if and only if the coefficients $\beta_1, \beta_2, \dots, \beta_n$, can be split into two groups with equal sums. Otherwise, the minimum is achieved at one of the given points a_i.

Remark 1. It seems that the conclusion of Theorem 3, as summarised in Bogomolny (2021), is not rigorous. For example, consider the function:

$$f(x) = \frac{(\pi-2)}{4}|x-1| + \frac{1}{3}|x-2| + \frac{(4-\pi)}{4}|x-3| + \frac{1}{6}|x-4|, \mathrm{x} \in \mathbb{R},$$

then $\beta_1 + \beta_3 = \beta_2 + \beta_4 = 1/2$. However, $x = 2$ is the unique minimiser of f.

Let $S = \sum_{j=1}^{n} \beta_j$ in (4). Since minimising f in (4) is the same as minimising f/S, without loss of generality, assume $S = 1$.

Theorem 3 only considers the sum of finite terms. One can also consider minimising functions of the form:

$$f(x) = \sum_{i=1}^{\infty} \beta_i|x-a_i|, \beta_i > 0, i = 1,2, \dots, \tag{5}$$

where $a_i, i = 1,2,$ are distinct. If $\sum_{i=1}^{\infty} \beta_i = 1$, thus $\beta_i, i = 1,2,$ is a *standardised sequence of coefficients*. The authors make this assumption in the following sections.

Besides being the sum of infinitely many terms, (5) is very different from (4). Since there are only finite terms in (4), one can always assume $a_1 < a_2 < \cdots < a_n$ by rearranging them. However, this operation may be not possible in (5). For example, assume $a_i = (-1)^i/\mathrm{i}, \mathrm{i} = 1,2, \dots$ One cannot rearrange the sequence to make it monotone; another difference is the set of minimisers. Theorem 3 indicates that at least one of the elements in $\{a_i: i = 1, \dots, n\}$ is a minimiser of f in (3), however this may be not true for f in (5). Example 3 in Section 3 shows a case where the unique minimiser of f is not an element in $\{a_i: i = 1,2, \dots\}$.

The purpose of this paper was to introduce a united probabilistic approach which can be used for minimising functions of the sum of finite or countably infinite terms of absolute values. The conclusion of Theorem 3 is a special case of the obtained result.

2. Theory

First, let us review the definition of *median* in probability theory.

Definition 1. Given a real-valued random variable Y, real number m is called a median of Y if

$$P\{Y \leq m\} \geq 1/2 \quad \text{and} \quad P\{Y \geq m\} \geq 1/2,$$

where $P(E)$ denotes the probability of event E. Given probability distribution function F, m is called a median of F if $F(m) \geq 1/2$ and $F(m-) \leq 1/2$.

This definition was given in Problem 1.7 in Lehmann and Casella (1998, p. 62). For example, suppose the distribution of Y is $P\{Y = 1\} = P\{Y = 2\} = 1/2$. Then any point in $[1,2]$ is a median of Y. Suppose the distribution of Y is $P\{Y = 1\} = 1/3$, $P\{Y = 2\} = P\{Y = 3\} = 1/4$, and $P\{Y = 5\} = 1/6$. Then Y has a unique median 2.

First, let us prove the existence of median(s) for any real-valued random variable.

Theorem 4. *Any real-valued random variable Y has at least one median, its set of medians is a finite closed interval. For any $a \in \mathbb{R}$, this paper regards the singleton $\{a\}$ as a closed interval $[a,a]$.*

Proof. Let $A = \{m \in \mathbb{R}: P\{Y \leq m\} \geq 1/2\}$ and $B = \{m \in \mathbb{R}: P\{Y \geq m\} \geq 1/2\}$. Then $A \cap B$ is the set of medians of Y. Since $\lim\limits_{m\to-\infty} P\{Y \geq m\} = \lim\limits_{m\to\infty} P\{Y \leq m\} = 1$, both A and B are nonempty. Let m_1 = inf A and m_2 = sup B. If $m_1 \notin A$, then $P\{Y \leq m_1\} < 1/2$, and there exists an $\varepsilon > 0$ such that $P\{Y \leq m_1 + \varepsilon\} < 1/2$, a contradiction. For any $\varepsilon > 0$, $P\{Y \geq m_2 - \varepsilon\} \geq 1/2$, which means that $P\{Y < m_2 - \varepsilon\} < 1/2$. Since $\varepsilon > 0$ is arbitrary, one has $P\{Y < m_2\} < 1/2$ and $P\{Y \geq m_2\} \geq 1/2$, which means $m_2 \in B$.

One can prove $m_1 \leq m_2$. Suppose it is not. Then $P\{Y \leq m_2\} < 1/2$ and there exists an $\varepsilon > 0$ such that $P\{Y \leq m_2 + \varepsilon\} < 1/2$, which means that $P\{Y > m_2 + \varepsilon\} > 1/2$, a contradiction.

It is clear that $m_1, m_2 \in A \cap B$, and if $m \in [m_1, m_2]$, then $m \in A \cap B$. If $m < m_1$, then $m \notin A$. If $m > m_2$, $m \notin B$, therefore $A \cap B = [m_1, m_2]$. QED.

Remark 2. This theorem is Part (b) of Problem 1.7 in Lehmann and Casella (1998, p. 62); another proof can be found in Shao (2005). Theorem 3 indicates that Y has either a unique median or uncountably infinitely many medians.

Theorem 5. *Given a random variable Y and $x \in \mathbb{R}$, let $E|Y - x|$ be the expectation of $|Y - x|$.*

(a) $E|Y - x| \geq E|Y - m|$ if m is a median of Y.

(b) If $E|Y| < \infty$ and x_0 minimises $f(x) = E|Y - x|$, $x \in \mathbb{R}$, then x_0 is a median of Y.

Proof. (a) If $E|Y| = +\infty$, then $E|Y - x| \geq E|Y| - x = +\infty$ for any $x \in \mathbb{R}$. The conclusion is trivial. Without loss of generality, assume $E|Y| < \infty$. Let m be any median of Y. First assume $x > m$. Then

(i) if $Y > x$,

$$|Y - x| - |Y - m| = -(x - m);$$

(ii) if $m < Y \le x$,

$$|Y - x| - |Y - m| = x + m - 2Y \ge x + m - 2x = -(x - m); \text{ and}$$

(iii) if $Y \le m$,

$$|Y - x| - |Y - m| = (x - m).$$

Therefore,

$$\begin{aligned} E|Y - x| - E|Y - m| &\ge (x - m)[P\{Y \le m\} - P\{Y > m\}] \\ &= (x - m)[2P\{Y \le m\} - 1] \ge 0. \end{aligned}$$

Similarly, if $x < m$,

$$\begin{aligned} &E|Y - x| - E|Y - m| \ge (m - x)[P\{Y \ge m\} - P\{Y < m\}] \\ &= (m - x)[2P\{Y \ge m\} - 1] \ge 0. \end{aligned}$$

Hence, $E|Y - x| \ge E|Y - m|$ for any $x \in \mathbb{R}$.

(b) Let $[m_1, m_2]$ be the set of medians of Y. Suppose $x_0 > m_2$. Then $P\{Y \ge x_0\} < 1/2$, and

$$E|Y - x_0| - E|Y - m_2| = (x_0 - m_2)[2P\{Y \le m_2\} - 1] + 2E[(x_0 - Y)\mathbf{1}_{\{m_2 < Y < x_0\}}].$$

If $P\{Y \le m_2\} > 1/2$, then $E|Y - x_0| - E|Y - m_2| > 0$. Let $x_1 = (m_2 + x_0)/2$. If $P\{Y \le m_2\} = 1/2$, then

$$\begin{aligned} E|Y - x_0| - E|Y - m_2| &= 2E\left[(x_0 - Y)\mathbf{1}_{m_2 < Y < x_0}\right] \\ &\ge 2E\left[(x_0 - Y)\mathbf{1}_{m_2 < Y < x_1}\right] \\ &\ge 2(x_0 - x_1)P\{m_2 < Y < x_1\} \\ &\ge 2(x_0 - x_1)[P\{Y < x_1\} - P\{Y > m_2\}] \\ &> 0. \end{aligned}$$

Similarly, one has $E|Y - x_0| - E|Y - m_1| > 0$ if $x_0 < m_1$. QED.

Remark 3. Part (a) of Theorem 5 is Problem 1.8 in Lehmann and Casella (1998, p. 62); another proof can be found in Shao (2005). This theorem means that if $E|Y| < \infty$, then x_0 minimises function $f(x) = E|Y - x|$ f and only if x_0 is a median of Y.

3. Applications

Theorem 5 offers a general united approach to finding minimum of functions in (4) and (5). Let Y be a random variable with probability distribution $P\{Y = a_i\} = \beta_i, i = 1,2, \ldots$. Then f in (4) or (5) is the same as

$$f(x) = E|Y - x|.$$

Since $f(x) = \infty$ if $E|Y| = \infty$, only consider the case that $E|Y| < \infty$, i.e. $\sum_{i=1}^{\infty} \beta_i |a_i| < \infty$. From Theorem 5 it follows that the key step in minimising f in (5) is to find a median of Y. For this purpose, the authors considered several cases:

Case 1: Assume $a_i, i = 1,2, \ldots$, is a strictly monotone sequence.

(i) Assume $a_1 < a_2 < \cdots$. For any $x \in \mathbb{R}$,

$$P\{Y \le x\} = \sum_{i: a_i \le x} \beta_i \quad \text{and} \quad P\{Y \ge x\} = \sum_{i: a_i \ge x} \beta_i.$$

Let

$$k_l = \inf\left\{ j \geq 1 : \sum_{i=1}^{j} \beta_i \geq 1/2 \right\} \quad \text{and} \quad k_u = \sup\left\{ j \geq 1 : \sum_{i=j}^{\infty} \beta_i \geq 1/2 \right\}$$

Since

$$\sum_{i=k_l}^{\infty} \beta_i = 1 - \sum_{i=1}^{k_l-1} \beta_i > 1 - 1/2 = 1/2,$$

Hence

$$k_l \leq k_u.$$

If $\sum_{i=1}^{k_l} \beta_i > 1/2$ then $\sum_{i=k_l+1}^{\infty} \beta_i < 1/2$, which means $k_u < k_l + 1$. This implies $k_u = k_l$, and a_{k_l} is the unique median of Y. If $\sum_{i=1}^{k_l} \beta_i = 1/2$, then $\sum_{i=k_l+1}^{\infty} \beta_i = 1/2$, which means $k_u = k_l + 1$, and the set of medians of Y is $[a_{k_l}, a_{k_u}]$.

In this case, at least one of the elements in $\{a_i : i = 1,2, \dots\}$ is a median of Y. Theorem 3 is a special case of this result.

Example 1. Consider f in (1). The standardised sequence of coefficients is

$$\frac{1}{8}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{4}.$$

It is clear that $k_l = 3$ and $k_u = 4$, therefore, f is minimised when $x = a_3 = 2$.

Example 2. Let

$$f(x) = \frac{1}{4}|x-1| + \frac{3}{4}\sum_{i=2}^{\infty} 2^{-i+1}|x-i|, \;\; x \in \mathbb{R}.$$

Then $k_l = k_u = 2$ and 2 is the unique median of Y. The minimum of f is 1.

(ii) Assume $a_1 > a_2 > \cdots$. Let $Y' = -Y$, then

$$f(x) = E|Y - x| = E|-Y - (-x)| = E\left|Y' - (-x)\right|.$$

One can first find any median m of Y' using the method above, then $-m$ minimises f. At least one of the elements in $\{a_i : i = 1,2, \dots\}$ is a median of Y.

Case 2: Assume $|a_i|, i = 1,2, \dots$, is a monotone sequence.

(i) Assume $|a_1| \leq |a_2| \leq \cdots$. Let

$$S_1 = \sum_{i:\, a_i \leq 0} \beta_i \quad \text{and} \quad S_2 = \sum_{i:\, a_i > 0} \beta_i.$$

If $S_1 \geq 1/2$, let

$$k_l = \sup\left\{ j : a_j < 0, \sum_{i:\, i \geq j,\, a_i \leq 0} \beta_i \geq \frac{1}{2} \right\}.$$

Then a_{k_l} is a median of Y.

If $S_1 < 1/2$, let

$$k_l = \inf\left\{ j: a_j \geq 0,\, S_1 + \sum_{i:\, i\leq j,\, a_i>0} \beta_i \geq \frac{1}{2} \right\}.$$

Then a_{k_l} is a median of Y.

(ii) Assume $|a_1| \geq |a_2| \geq \cdots$. Using a similar idea, one can easily find a median of Y in $\{a_i: i = 1,2,\dots\}$.

In these two cases, at least one of the elements in $\{a_i: i = 1,2,\dots\}$ is a median of Y which minimises f. However, this may not be true in general cases.

Case 3: Assume $a_i, i = 1,2,\dots$, is a general sequence.

There are two challenges in this case: (i) generally there is no explicit way to find a median of Y, and (ii) even if the median of Y is unique, it may not be any element in $\{a_i: i = 1,2,\dots\}$. The authors used the following example to show these challenges.

Example 3. Let $a_1, a_2, \dots$, be an enumeration of rationals in $(0,1)$. Let f be in the form of (5). It was first proved that Y has a unique median m. Suppose is not, then let $m_1 < m_2$ be two medians of Y. Since rationals are dense in $(0,1)$,

$$\frac{1}{2} \leq P\{Y \leq m_1\} < P\{Y < m_2\} = 1 - P\{Y \geq m_2\} \leq \frac{1}{2},$$

a contradiction.

The median of Y may be rational or irrational.

(i) Suppose a is a rational in $(0,1)$. Let $\beta^{(a)} = P\{Y = a\}$. If $\beta^{(a)} \geq 1/2$, then a is the median of Y. If $\beta^{(a)} < 1/2$, then a is the median of Y if and only if

$$\sum_{i:a_i<a} \beta_i \geq \frac{1}{2} - \beta^{(a)} \quad \text{and} \quad \sum_{i:a_i>a} \beta_i \geq \frac{1}{2} - \beta^{(a)}.$$

(ii) Suppose a is an irrational in $(0,1)$. Then a is the median of Y if and only if

$$\sum_{i:a_i<a} \beta_i = \sum_{i:a_i>a} \beta_i = \frac{1}{2}. \tag{6}$$

Note that (6) holds only if a is irrational. In this case, although Y is a rational-valued random variable, its unique median is irrational.

Here is a numeric example. Let $s_i, i = 1,2,\dots$ be an enumeration of rationals in $(0, \sqrt{2}/2)$ and $t_i, i = 1,2,\dots$ be an enumeration of rationals in $(\sqrt{2}/2, 1)$. Let X be a random variable with $P\{X = s_i\} = P\{X = t_i\} = \frac{1}{2^{i+1}}, i = 1,2,\dots$. The distribution function F of X is strictly increasing on $[0,1]$, continuous at irrationals, and $F(\sqrt{2}/2) = 1/2$. Then $\sqrt{2}/2$ is the unique median of X, even though $P\{X = \sqrt{2}/2\} = 0$.

4. Discussion and Conclusions

In this paper, the authors presented a unified probabilistic approach that can be used to determine the minimum value of a broad class of functions represented by equation (5), where the sequence $a_i, , i = 1,2, \ldots$ is appropriately ordered. However, as discussed in Case 3 in Section 3, if the sequence is not ordered, and this method may not be applicable. Research on addressing this issue is currently underway.

Quantile regression was first introduced by Koenker & Bassett (1978) in the field of econometrics. One of the most notable cases of quantile regression is the *median regression* estimator, which minimises the sum of absolute errors. This study provides a theoretical justification for this method.

References

Bogomolny, A. (2021, March 20). *Sum of Absolute Values*. Cut the Knot. https://www.cut-the-knot.org/m/Algebra/MinimumWithAbsoluteValue.shtml

Koenker, R., & Bassett, G. (1978). Regression Quantiles. *Econometrica*, *46*, 33–50.

Lehmann, E. L. & Casella, G. (1998). *Theory of Point Estimation*. 2nd edition. Springer.

Nahin, P. J. (2004). *When Least Is Best*. Princeton University Press.

Shao, J. (2005). *Mathematical Statistics: Exercises and Solutions*. Springer.

Xu, J. (2012). *Lecture Notes on Mathematical Olympiad Courses*. For senior section, vol. 1. World Scientific.

Zunifikowane probabilistyczne podejście do minimalizacji sumy wartości bezwzględnych

Streszczenie

Cel: Propozycja nowej metody minimalizacji funkcji sum wartości bezwzględnych.

Metodyka: Suma wartości bezwzględnych jest standaryzowana w taki sposób, aby suma parametrów była równa 1. Wówczas suma wartości bezwzględnych przyjmuje wartości takie, że $E|X - a|$, gdzie X jest zmienną losową.

Wyniki: Dowolna mediana X minimalizuje funkcję $E|X - a|$. Aby zminimalizować funkcję, wystarczy znaleźć dowolną medianę X.

Implikacje i rekomendacje: Metoda przedstawiona w artykule ma zastosowanie do minimalizacji szerokiej rodziny funkcji.

Oryginalność/wartość: W naszej pracy wykorzystano metodę probabilistyczną do rozwiązania problemów optymalizacyjnych.

Słowa kluczowe: mediana, rozkład prawdopodobieństwa, wartość minimalna