TOWARDS A WELL–DEFINED MEDIAN

AHMAD AL-SALMAN AND MOWAFFAQ HAJJA

(communicated by J. Pečarić)

Abstract. The diagonal $\Delta$ of $\mathbb{R}^n$ is Chebyshev with respect to the $p$-norm for every $p \in (1, \infty]$ but not for $p = 1$. As a result, the median is multi-valued, since the median of a data set $\{a_1, \ldots, a_n\}$ can be thought of as the number(s) $\mu$ for which the point $(\mu, \ldots, \mu)$ is a point on $\Delta$ that best approximates the point $(a_1, \ldots, a_n)$ with respect to the $\ell_1$-norm. In this note, it is proved that if $(\mu_p, \ldots, \mu_p)$ is the unique point on $\Delta$ that best approximates a fixed point $(a_1, \ldots, a_n)$ with respect to the $\ell_p$-norm for $p \in (1, \infty]$, then as $p$ decreases to 1, $\mu_p$ converges, and its limit is proposed to be called the median of $\{a_1, \ldots, a_n\}$. Along the way, $\mu_p$ is shown to be continuous in $p$ for all $p \in (1, \infty]$ in the sense that $\mu_p$ converges to $\mu_q$ as $p$ goes to $q$ for every $q \in (0, \infty]$.

1. Preliminaries

We fix $n \geq 1$, and we let $\Delta$ stand for the diagonal of $\mathbb{R}^n$. Thus

$$\Delta = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n : x_1 = \cdots = x_n\}.$$ 

For $p \in [1, \infty]$, the $\ell_p$-norm on $\mathbb{R}^n$ is defined by

$$\|x\|_p = \begin{cases} \left(\sum_{j=1}^n |x_j|^p\right)^{1/p}, & \text{if } p \in [1, \infty) \\ \max\{|x_j| : j = 1, \ldots, n\}, & \text{if } p = \infty, \end{cases}$$

where $x = (x_1, \ldots, x_n)$. Note that $\|x\|_\infty = \lim_{p \to \infty} \|x\|_p$.

It is easy to see that $\Delta$ is Chebyshev in $\mathbb{R}^n$ with respect to the $\ell_p$-norm for all $p \in (1, \infty]$. In other words, given any $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ and any $p \in (1, \infty]$, there is a unique point $(\mu_p, \ldots, \mu_p)$ in $\Delta$ whose $\ell_p$-distance from $a$ is minimal. This must be well-known, but it also follows from the discussion below. We keep $a = (a_1, \ldots, a_n)$ fixed throughout, and we let $(\mu_p, \ldots, \mu_p)$ be its best approximant in $\Delta$ with respect to the $\ell_p$-norm.

It is clear that $\mu_\infty$ is the mid-point of the range of the data set $\{a_1, \ldots, a_n\}$, i.e.,

$$\mu_\infty = \frac{\max\{a_1, \ldots, a_n\} + \min\{a_1, \ldots, a_n\}}{2}. \quad (1)$$


Key words and phrases: Median, mean, best approximation, convex, Chebyshev set, Fermat-Torricelli point.
Here, a data set is any set with possible repetitions. If \( p \in (1, \infty) \), then \( \mu_p \) is the unique point \( x \) for which the function \( f_p : \mathbb{R}^n \to \mathbb{R} \) defined by

\[
f_p(x) = \| (x, \cdots, x) - (a_1, \cdots, a_n) \|^p = \sum_{j=1}^{n} |x - a_j|^p
\]

attains its minimum. In other words, \( \mu_p \) is the unique zero of

\[
f'_p(x) = p \sum_{j=1}^{n} \text{sign}(x - a_j)|x - a_j|^{p-1},
\]

where \( \text{sign}(t) \) is defined to be 1 if \( t \geq 0 \) and \(-1\) otherwise. The existence and uniqueness of \( \mu_p \) follow from the strict convexity of \( f_p \) which in turn follows from the fact that

\[
f''_p(x) = p(p-1) \sum_{j=1}^{n} |x - a_j|^{p-2} > 0.
\]

The number \( \mu_p \) may be called the \( \ell_p \)-best-approximation mean of \( a \), or rather of the corresponding data set \( \{a_1, \cdots, a_n\} \). It is clear that \( \mu_p \) is strictly internal in the sense that

\[
\min\{a_1, \cdots, a_n\} \leq \mu_p \leq \max\{a_1, \cdots, a_n\},
\]

with strict inequality except in the case when \( a_1 = \cdots = a_n \). In fact, if \( a_1 \leq \cdots \leq a_n \) and if they are not all equal, then (3) shows that \( f'_p(a_1) < 0 \) and \( f'_p(a_n) > 0 \).

2. Continuity of the \( \ell_p \)-best-approximation mean for \( p \in (1, \infty) \)

In this section, we establish the continuity of \( \mu_p \) for \( p \in (1, \infty) \).

**Theorem 1.** For a fixed point \( a = (a_1, \cdots, a_n) \), let \( \mu_p, \ 1 < p < \infty \), be the unique value that minimizes the function \( f \) defined in (2), and let \( \mu_\infty \) be defined by (1). Then \( \mu_p \) is continuous for \( p \in (1, \infty) \). In other words,

\[
\lim_{p \to q} \mu_p = \mu_q \ \forall \ q \in (1, \infty].
\]

**Proof.** Let \( q \in (1, \infty] \) be given, and let \( p_k \) be any sequence that converges to \( q \) such that \( \mu_{p_k} \) converges to \( t \), necessarily finite since \( \mu_p \) is internal. We are to show that \( t = \mu_q \). For simplicity, set \( t_k = \mu_{p_k} \). In view of (3), \( \mu_p \) is defined by

\[
\sum_{j=1}^{n} \text{sign}(\mu_p - a_j)|\mu_p - a_j|^{p-1} = 0.
\]

Then

\[
\sum_{j=1}^{n} \text{sign}(t_k - a_j)|t_k - a_j|^{p_k-1} = 0. \tag{4}
\]
If \( q < \infty \), we take limits as \( k \) goes to infinity and we obtain
\[
\sum_{j=1}^{n} \text{sign}(t - a_j)|t - a_j|^{q-1} = 0.
\]

Therefore \( t \) satisfies the equation that defines \( \mu_q \), and therefore \( t = \mu_q \), as desired.

If \( q = \infty \), we shall show that \( t \) is equal to \( \mu_\infty \). Thus we assume that \( t \neq \mu_\infty \) and seek a contradiction. Without loss of generality, we assume that \( a_1 \leq \cdots \leq a_n \) and that \( t > \mu_\infty \). Then \( \mu_\infty = (a_n + a_1)/2 \), and our assumption takes the form \( t > (a_1 + a_n)/2 \).

Dividing (4) by the mid-range \( u = (a_n - a_1)/2 \) and taking the limit, we see that
\[
\lim_{k \to \infty} \left( \sum_{j=1}^{n} \text{sign}(t_k - a_j)|t_k - a_j|^{p_k-1} \right) = 0.
\]

(5)

Since \( t > (a_1 + an)/2 \), then every \( a_j \) for which \( t - a_j \) is negative has the property that \( |t - a_j| \) is less than \( |u| \). For such an \( a_j \),
\[
\lim_{k \to \infty} \left| \frac{t_k - a_j}{u} \right|^{p_k-1} = 0.
\]

Therefore we may discard the terms in (5) for which \( t - a_j < 0 \) or equivalently neglect their signs. We obtain
\[
\lim_{n \to \infty} \left( \sum_{j=1}^{n} \left| \frac{t_n - a_j}{u} \right|^{p_n-1} \right) = 0.
\]

This contradicts the assumption that \( |t - a_1| > |u| \) which implies that
\[
\lim_{k \to \infty} \left| \frac{t_k - a_1}{u} \right|^{p_k-1} = \infty.
\]

Therefore \( t = (a_1 + a_n)/2 \), as desired. \( \square \)

3. Towards a unique median

We now turn to the case \( p = 1 \). Since the function \( f_1 \) defined in (2) attains its minimum but not necessarily at a unique point, it follows that \( \Delta \) is proximinal but not Chebeshev with respect to the \( \ell_1 \)-norm. In fact, if \( a_1 \leq \cdots \leq a_n \), then \( f_1 \) attains its minimum at the unique point \( a_{(n+1)/2} \) if \( n \) is odd and at any point in the interval \([a_{n/2}, a_{1+n/2}]\) if \( n \) is even. This explains the non-uniqueness of the median as ordinarily defined in statistics.

In this section, \( a = (a_1, \cdots, a_n) \) will stand for an arbitrary, but fixed, point in \( \mathbb{R}^n \), and \( \mu_p = \mu_p(a) \), \( p \in (1, \infty] \), is defined, as above, to be the point at which the function \( f_p \) defined in (2) attains its minimum. We shall show that the limit of \( \mu_p \) as \( p \) decreases to 1 exists, and we propose to call this limit, that we denote by \( \mu^* \)
median of \(a\), or rather of the data set \(\{a_1, \ldots, a_n\}\). It is obvious that if \(n=2\), then 
\[\mu_p = (a_1 + a_n)/2\] for all \(p\). Also, if all the \(a_j\) are equal, then \(\mu_p\) equals their common value for all \(p\). Thus we exclude the trivial cases \(n<3\) and \(a_1 = \cdots = a_n\). Also, there is no loss in generality in assuming that \(a_1 \leq \cdots \leq a_n\).

It is expectedly impossible to have a closed formula for \(\mu_p\). However, given \(w \in \mathbb{R}\), it is easy to decide whether \(w = \mu_p\), \(w < \mu_p\), or \(w > \mu_p\). In fact, having the parabola-like graph of \(f_p\) in mind, it is obvious that
\[
w < \mu_p, \quad w = \mu_p, \quad \text{or} \quad w > \mu_p\]
accordingly as
\[
f'_p(w) < 0, \quad f'_p(w) > 0, \quad \text{or} \quad f'_p(w) = 0,\]
respectively. \((6)\)

We shall freely use this simple, but quite useful, observation.

We start with a simple lemma that will be used in proving the key Lemma 3. This in turn will be used in proving our main Theorem 4.

**Lemma 2.** Let \(\alpha_1, \ldots, \alpha_n, \ n \geq 1\), be non-zero real numbers, and let \(r_1, \ldots, r_n\) be distinct real numbers. Then the function
\[F(x) = \alpha_1 e^{r_1x} + \cdots + \alpha_n e^{r_nx}\]
has finitely many zeros in \(\mathbb{R}\). The same holds for the function
\[\alpha_1 c_1^x + \cdots + \alpha_n c_n^x,\]
where \(c_1, \ldots, c_n\) are any distinct positive real numbers.

**Proof.** Without loss of generality, assume that \(r_1 < \cdots < r_n\), and let
\[g(x) = e^{-r_1x}f(x) \quad \text{and} \quad h(x) = e^{-r_nx}f(x).\]
Then
\[
\lim_{x \to -\infty} g(x) = \alpha_1 \quad \text{and} \quad \lim_{x \to +\infty} h(x) = \alpha_n.
\]
Therefore \(g\) has no zeros near \(-\infty\) and \(h\) has no zeros near \(\infty\). Since \(F\), \(g\), and \(h\) have the same zeros, it follows that the zeros of \(F\) all lie in some compact interval. If the set of zeros of \(F\) is infinite, it has an accumulation point and \(F\) would have to be identically zero. This follows because \(F\), as a function of the complex variable \(x\), is entire. Therefore \(F\) has finitely many zeros, as claimed.

The last statement follows by setting \(c_j = e^{r_j}\). \(\square\)

**Lemma 3.** For \(p > 1\), let \(f_p : \mathbb{R} \to \mathbb{R}\) be defined as in \((2)\) and let \(\mu_p\) be the unique point at which \(f_p(x)\) attains its minimum. Let \(w\) be any real number. Then there exists \(\delta > 0\) such that either \(\mu_p \geq w\) for all \(p \in (1, 1+\delta)\), or \(\mu_p \leq w\) for all \(p \in (1, 1+\delta)\).

**Proof.** We suppose, by way of contradiction, that there are sequences \(p_i\) and \(q_i\) decreasing to 1 such that \(\mu_{q_i} < w < \mu_{q_i}\). It follows from \((6)\) that
\[
f'_{q_i}(w) > 0 > f'_{p_i}(w).\] \((7)\)
Now let
\[ F(p) = \frac{f_p'(w)}{p} = \sum_{j=1}^{n} \text{sign}(w - a_j)|w - a_j|^p-1. \]

Then (7) says that \( F(q_i) > 0 > F(p_i) \). Since \( F \) is continuous as a function of \( p \), it follows from the intermediate value theorem that \( F(\xi_i) = 0 \) for some \( \xi_i \) between \( p_i \) and \( q_i \). Also \( \xi_i \) decreases to 1 by the Sandwich theorem. Therefore 1 is an accumulation point of the set of zeros of \( F \), contradicting Lemma 2 and completing the proof. \( \square \)

**Theorem 4.** For \( p > 1 \), let \( f_p : \mathbb{R} \to \mathbb{R} \) be defined as in (2) and let \( \mu_p \) be the unique point at which \( f_p(x) \) attains its minimum. Then as \( p \) decreases to 1, \( \mu_p \) converges, say to \( \mu^* \). Also, the convergence is ultimately one-sided in the sense that there exists \( \delta > 0 \) such that \( \mu_p - \mu^* \) has the same sign for all \( p \in (1, 1 + \delta) \). If the number of indices \( i \) for which \( a_i \) is greater than \( \mu^* \) is \( M \) and the number of indices \( i \) for which \( a_i \) is less than \( \mu^* \) is \( m \), then \( M \leq n/2 \) and \( m \leq n/2 \). Also, if \( M > m \), then the convergence is ultimately right-sided, and if \( M < m \), then the convergence is ultimately left-sided. Consequently \( \mu^* \) lies in the traditional median set of \( a \) in the sense that \( \mu^* \) equals \( a_{(n+1)/2} \) if \( n \) is odd and lies in the interval \( [a_{n/2}, a_{1+n/2}] \) if \( n \) is even.

**Proof.** If \( \lim sup_{p \to 1} \mu_p \) and \( \lim inf_{p \to 1} \mu_p \) are not equal, then any real number \( w \) between them would contradict Lemma 3. Therefore \( \lim_{p \to 1} \mu_p \) exists. Let us denote it by \( \mu^* \). Taking \( w = \mu^* \) in Lemma 3, we conclude that the convergence is ultimately one-sided.

Now let us treat the case when the convergence of \( \mu_p \) is ultimately right-sided. Thus \( \mu_p \geq \mu^* \) for all \( p \) close enough to 1. Let \( m, s \), and \( M \) be defined by
\[
a_1 \leq \cdots \leq a_m < a_{m+1} = \cdots = a_{m+s} = \mu^* < a_{m+s+1} \leq \cdots \leq a_{m+s+M},
\]
where \( m, M \geq 0 \), \( s \geq 1 \), and \( m + s + M = n \). Let \( a \) be the common value of \( a_{m+1}, \cdots, a_{m+s} \). Then in a right neighborhood of \( \mu^* \), we have
\[
f_p'(x) = p \sum_{j=1}^{m} (x - a_j)^{p-1} + ps(x - a)^{p-1} - p \sum_{j=m+s+1}^{m+s+M} (a_j - x)^{p-1}.
\]

Therefore,
\[
s(\mu_p - a)^{p-1} = \sum_{j=m+s+1}^{m+s+M} (a_j - \mu_p)^{p-1} - \sum_{j=1}^{m} (\mu_p - a_j)^{p-1}.
\]

The right-hand side converges, as \( p \) decreases to 1 (and consequently \( \mu_p \) decreases to \( \mu^* \)) to \( M - m \). Therefore (the indeterminate form) \( (\mu_p - a)^{p-1} \) converges, as \( p \) decreases to 1, to some \( L \) with \( sL = M - m \). Also, \( L \in [0, 1] \), because \( (\mu_p - a)^{p-1} \in [0, 1] \) for all \( p \) sufficiently close to 1. Therefore \( 0 \leq M - m \leq s \). Since \( s = n - m - M \), it follows that \( M - m \leq n - m - M \), and therefore \( M \leq n/2 \). Hence \( m \leq M \leq n/2 \),
as claimed. This also shows that if \( m > M \), then the convergence cannot be ultimately right-sided and therefore it must be left-sided.

The case when the convergence of \( \mu_p \) is ultimately left-sided is similar, and the proof is complete. □

The ultimately left-sidedness of the convergence of \( \mu_p \) raises the question whether the convergence of \( \mu_p \) is indeed ultimately monotone. Thus given \( a_1, \ldots, a_n \) and letting \( \mu_p \) be as defined earlier, does there exist \( \delta > 0 \) such that the convergence of \( \mu_p \) as \( p \) decreases to 1 within the interval \((1, 1 + \delta)\) is monotone. The next theorem provides an affirmative answer when \( n = 3 \).

**Theorem 5.** If \( n = 3 \), then \( \mu_p \) converges monotonically to the median.

**Proof.** Let \( a \leq b \leq c \) be given. We are to show that \( \mu_p = \mu_p(a, b, c) \) converges ultimately monotonically. If \( b - a = c - b \), then it follows from symmetry that \( \mu_p = b \) for all \( p \). Therefore assume that \( b - a > c - b \).

Recall that \( \mu_p \) is the unique point at which \( f_p(x) = |x - a|^p + |x - b|^p + |x - c|^p \) attains its minimum. Since

\[
f'_p(b) = p(b - a)^{p-1} - p(c - b)^{p-1} > 0,
\]

it follows from (6) that \( \mu_p < b \), and therefore \( \mu \) lies in the open interval \((a, b)\). For \( x \in (a, b) \),

\[
f_p(x) = (x - a)^p + (b - x)^p + (c - x)^p,
\]

and

\[
f'_p(x) = p(x - a)^{p-1} - p(b - x)^{p-1} - p(c - x)^{p-1}.
\]

Therefore

\[
(\mu_p - a)^{p-1} - (b - \mu_p)^{p-1} - p(c - \mu_p)^{p-1} = 0.
\]

Let \( q > p \). To show that \( \mu_q < \mu_p \), it is enough to show that \( f'_q(\mu_p) < 0 \). In other words, we are to show that \( (\mu_p - a)^{q-1} - (b - \mu_p)^{q-1} - (c - \mu_p)^{q-1} < 0 \).

Let

\[
(\mu_p - a)^{p-1} = \alpha, (b - \mu_p)^{p-1} = \beta, (c - \mu_p)^{p-1} = \gamma, \quad \frac{q-1}{p-1} = s > 1.
\]

Then \( \alpha, \beta, \) and \( \gamma \) are positive, and \( \alpha = \beta + \gamma \). Therefore

\[
\alpha^s = (\beta + \gamma)^s > \beta^s + \gamma^s,
\]

or equivalently

\[
(\mu_p - a)^{q-1} - (b - \mu_p)^{q-1} - (c - \mu_p)^{q-1} > 0.
\]

Thus \( f'_q(\mu_p) > 0 \), and therefore \( \mu_q < \mu_p \).

Therefore as \( p \) decreases to 1, \( \mu \) increases, necessarily to \( b \). □

The question of monotonicity of convergence when \( n > 3 \) remains open.
4. Relation to distance means and genuine uniqueness of $\mu^*$

It is useful to view the afore-mentioned best approximation means $\mu_p$ from the perspective of the distance means introduced in [3]. Thus we start with any distance $d$ on a set $X$ for which the function

$$f_d(x) = d(x, a_1) + \cdots + d(x, a_n)$$

attains its minimum at a unique point $v = v_d(a)$ for every data set $a = \{a_1, \cdots, a_n\}$ in $X$, and we call the point $v_d(a)$ the distance mean of $a$ associated with $d$. Obviously, the $\ell_p$-best-approximation mean $\mu_p$, $p \in (1, \infty)$, is nothing but the distance mean associated with the distance $d_p$ defined on $\mathbb{R}$ by

$$d_p(x, y) = |x - y|^p.$$  \hspace{1cm} (8)

Here, a distance on $X$ is a real-valued symmetric positive-definite function on $X \times X$ that does not necessarily satisfy the triangle inequality; see [2]. Note that $\mu_\infty$ does not correspond, at least not obviously, to a distance.

This viewpoint has the advantage of allowing one to talk simultaneously about means of sets of different sizes. More importantly, it allows us to think of several triangle (and generally simplex) centers as distance means. For example, considering the Euclidean norm $\|\cdots\|$ on $\mathbb{R}^n$ and letting $\alpha$ be the distance defined on $\mathbb{R}^n$ by

$$\alpha(x, y) = \|x - y\|^2, \ x, y \in \mathbb{R}^n,$$

we immediately see that $v_\alpha(x_1, \cdots, x_N)$ is nothing but the centroid of the points $x_1, \cdots, x_N$ for every $x_1, \cdots, x_N \in \mathbb{R}^n$. Letting $\beta$ be defined by

$$\beta(x, y) = \|x - y\|, \ x, y \in \mathbb{R}^n,$$

we see that $v_\beta(x_1, \cdots, x_N)$ exists and is unique only if the points $x_1, \cdots, x_N$ are non-collinear, and that in this case it is nothing but the Fermat-Torricelli point of these points; see [4]. To fit these considerations in the context of best approximation means, one would have to work with unusual spaces of the form $(\mathbb{R}^N)^n$ with

$$\|(x_1, \cdots, x_n)\|_p = \left(\sum_{j=1}^n \|x_j\|_p^p\right)^{1/p}, \ x_j \in \mathbb{R}^N.$$

The distance-mean approach also raises a question regarding the uniqueness of the path that we have taken in defining $\mu^*$. More precisely, it is conceivable that there exist, beside the distances $d_p$ defined in (8), other sequences $\delta_p$ of distances on $\mathbb{R}$ that converge, as $p$ decreases to 1, to the ordinary Euclidean distance and such that the limit $\mu^\dagger$ of their associated distance means is not the same as our $\mu^*$. In this case, $\mu^\dagger$ would have as valid a claim as $\mu^*$ to being the median. However, the fact that the distances $d_p$ that we have chosen are highly natural may help us accept $\mu^*$ as a highly legitimate median.
Another approach for defining the median of a data set \( a = \{a_1, \cdots, a_n\} \) can be based on the Fermat-Torricelli point described above. As noted then, the function defined on \( \mathbb{R}^N \) by
\[
F(x) = \|x - x_1\| + \cdots + \|x - x_n\| \tag{9}
\]
attains its minimum at a unique point, the so-called Fermat-Torricelli point, if and only if \( x_1, \cdots, x_n \) are non-collinear. This raises the possibility of defining the median of \( a_1, \cdots, a_n \) by embedding \( \mathbb{R} \) in \( \mathbb{R}^N \) for \( N \geq 2 \), perturbing the points \( a_i = (a_i, 0, \cdots, 0) \) in such a way that they stop being collinear, finding the Fermat-Torricelli point of the resulting set of points, and then taking the limit as the \( a_i \) move back to their initial collinear position. One wonders whether there are natural ways for performing the perturbation, and how the resulting limit is related to \( \mu^* \).

It is of course legitimate to question the usefulness of the proposed \( \mu^* \) and other possible definitions of the median, and that can be left to statisticians and workers in relevant disciplines to decide.

It would also be interesting to investigate whether \( \mu_p \), already proved to be continuous, is differentiable in \( p \), and whether each \( \mu_p \) is continuous or even differentiable in \( a \). One also raises the same question for \( \mu^* \).

Finally, one may consider continuous versions of the above investigation by looking at spaces of integrable functions on measure spaces. In this case, one would restrict attention to spaces of finite measure to ensure that diagonal, i.e., constant, functions are included.

\textit{Acknowledgments.} The authors are supported by research grants from Yarmouk University, and are grateful for this support. The second-named author is also grateful to Mr. Zuhair Abu-Abbas for fruitful discussions, several years ago, that resulted in another proof that \( \mu_p \to \mu_\infty \) as \( p \to \infty \) \cite{1}.

\textbf{REFERENCES}

\begin{enumerate}
\item Z. ABU-ABBAS, Private Communications, 1995.
\item M. HAJJA, Distance means and best approximation means: Some elementary questions, preprint.
\end{enumerate}

Ahmad Al-Salman  
Mathematics Department  
Yarmouk University  
Irbid, Jordan  
e-mail: alsalman@yu.edu.jo

Mowaffaq Hajja  
Mathematics Department  
Yarmouk University  
Irbid, Jordan  
e-mail: mhajja@yu.edu.jo  
e-mail: mowhajja@yahoo.com

Journal of Mathematical Inequalities  
www.ele-math.com  
jmi@ele-math.com

(Received August 27, 2006)