Properties of ordinary and weighted sums of order $p$ used for distance estimation


<http://www.numdam.org/item?id=RO_1995__29_1_59_0>
PROPERTIES OF ORDINARY AND WEIGHTED SUMS OF ORDER \( p \) USED FOR DISTANCE ESTIMATION (*)

by J. BRIMBERG (1) and R. F. LOVE (2)

Communicated by Brian BOFFEY

Abstract. - Sums of order \( p \) are encountered in numerous practical applications; an example is the popular \( l_p \) norm found in continuous location theory. This paper attempts to generalize various properties of ordinary and weighted sums of order \( p \) which are of theoretical and practical interest. We expand the conditions for which Jensen's inequality holds. In addition, an important convexity result is generalized and an open question posed by Beckenbach is resolved.

Keywords: Weighted sums, order \( p \), distance estimation.

1. INTRODUCTION


(*) Received September 1993.
(1) Department of Engineering Management, Royal Military College of Canada, Kingston, Ontario, K7K 5L0, Canada.
(2) Department of Management Science/Systems, McMaster University, Hamilton, Ontario, L8S 4M4, Canada.
construction of Voronoi diagrams. The use of p.d.f.s in location-allocation models is given by Love et al. [13]. Two of the most widely used truck scheduling software packages in North America, Truckstops [16] and Roadnet [15] incorporate p.d.f.s since they are much more efficient and comprehensive to use in practice than attempting to assemble large files of distance data.

Love and Morris [10, 11, 12] applied several distance norms including the weighted $l_p$ ($kl_p$) norm to Germany and several regions of the United States. Let $x = (x_1, x_2)^T$, $y = (y_1, y_2)^T$ be any two points in the plane. The $kl_p$ norm is given by

$$kl_p(x, y) = k[|x_1 - y_1|^p + |x_2 - y_2|^p]^{1/p}, \quad k > 0, \quad p > 0.$$  

Love and Morris found that the $kl_p$ norm was relatively easy to fit to a geographical region and it has excellent predictive properties. Ward and Wendell [17, 18] have introduced the concept of utilizing block norms as distance predictors. A new study by Love and Walker [14] shows that although marginally better results can be obtained by using a block norm with eight or more parameters than by using the $kl_p$ norm, the computation cost of fitting the block norms can be prohibitive. Conversely, the original studies by Love and Morris [10, 11] show that the $kl_p$ norm usually gives much superior results compared to other simpler norms such as the weighted Euclidean or weighted rectangular norms.

In the present paper we introduce a generalized $kl_p$ norm in the form of a weighted sum of order $p$. This is in effect adding a single parameter to the $kl_p$ norm since one of the two weights in the sum of order $p$ function replaces the $k$ in the $kl_p$ norm. Properties of the weighted sum as a function of the parameter $p$ are derived which can be utilized when fitting the norm to actual geographical data.

2. PRELIMINARY RESULTS

A weighted sum of order $p$ is defined as follow (e.g., section 2.10 of [6]):

$$T(y; b, p) = \left[ \sum_{i=1}^{K} b_i y_i^p \right]^{1/p}, \quad (1)$$
PROPERTIES OF ORDINARY AND WEIGHTED SUMS OF ORDER $p$

where

$$y = (y_1, ..., y_K)^T, \quad y_i > 0, \quad i = 1, ..., K,$$

$$b = (b_1, ..., b_K)^T, \quad b_i > 0, \quad i = 1, ..., K,$$

and $p \neq 0$.

The vector $b$ and the scalar $p$ can be considered as a set of parameter values. If all the weights $b_i = 1$, then $T$ becomes the ordinary sum of order $p$ which is well-known in the literature (e.g., section 1-16 of [2]). Note that the function $T(y; b, p)$ has the form of a generalized $l_p$ distance given by

$$l_{bp}(x) = \left[ \sum_{i=1}^{K} b_i |x_i|^p \right]^{1/p}, \quad (2)$$

where $x = (x_1, ..., x_K)^T \in \mathbb{R}^K$, $p$ is generally assumed to be greater than zero, and $l_{bp}$ estimates the distance between any two points $y, z \in \mathbb{R}^K$ such that $x = y - z$. The weights $b_i$ can be used to represent non-symmetric costs along the axis directions in a location model. (A comprehensive treatment of continuous location models is given in [13].) We are interested here in studying the behaviour of the sum $T$ as a function of its parameter $p$.

Consider first the asymptotic behaviour of $T$. Letting $y_m = \min_i (y_i)$ and $y_M = \max_i (y_i)$, we obtain

$$\lim_{p \to +\infty} \{ T(y; b, p) \} = \lim_{p \to +\infty} \left\{ \left[ \sum_{i=1}^{K} b_i y_i^p \right]^{1/p} \right\}$$

$$= y_M \lim_{p \to +\infty} \left\{ \left[ \sum_{i=1}^{K} b_i \left( \frac{y_i}{y_M} \right)^p \right]^{1/p} \right\}$$

$$= y_M, \quad (3)$$

and similarly,

$$\lim_{p \to -\infty} \{ T(y; b, p) \} = y_m \lim_{p \to -\infty} \left\{ \left[ \sum_{i=1}^{K} b_i \left( \frac{y_m}{y_i} \right)^{-p} \right]^{1/p} \right\}$$

$$= y_m. \quad (4)$$
Thus, the function $T$ approaches the same horizontal asymptotes irrespective of the positive weights $b_i, i = 1, \ldots, K$.

Without loss in generality, let us assume that all the $y_i$'s have distinct values; that is, $y_i \neq y_j, i \neq j$, for all $i, j \in \{1, \ldots, K\}$. (If this is not the case, common terms can be added together and $K$ adjusted accordingly.) Denoting the weights associated with $y_m$ and $y_M$ by $b_m$ and $b_M$ respectively, it is clear from (3) and (4) that for $K \geq 2$,

$$
\lim_{p \to +\infty} T = \begin{cases} 
    y_M^+, & \text{if } b_M \geq 1, \\
    y_M^-, & \text{if } 0 < b_M < 1,
\end{cases}
$$

and

$$
\lim_{p \to -\infty} T = \begin{cases} 
    y_m^+, & \text{if } b_m \geq 1, \\
    y_m^-, & \text{if } 0 < b_m < 1.
\end{cases}
$$

Thus, the direction of approach from above or below the horizontal asymptotes $y_M, y_m$ depends on the magnitude of the corresponding weights $b_M, b_m$.

We now examine the behaviour of $T$ near $p = 0$. Letting

$$
\beta = \sum_{i=1}^{K} b_i,
$$

it is readily seen that for $\beta > 1$,

$$
\lim_{p \to 0^+} T = +\infty, \quad \lim_{p \to 0^-} T \to 0; \quad (8)
$$

while for $\beta < 1$,

$$
\lim_{p \to 0^+} T = 0, \quad \lim_{p \to 0^-} T = +\infty. \quad (9)
$$

When $\beta = 1$, $T(y; b, p)$ becomes the mean value function, for which the following well-known result applies:

$$
\lim_{p \to 0} T = \prod_{i=1}^{K} y_i^{b_i}. \quad (10)
$$

It follows from (8), (9) and (10) that $T$ is continuous at $p = 0$ if, and only if, $\beta = 1$. 

Recherche opérationnelle/Operations Research
We now calculate the first and second-order partial derivatives of \( \ln T \) with respect to \( p \). Letting

\[
a_i = \frac{b_i}{\beta}, \quad i = 1, \ldots, K,
\]

equation (1) can be rewritten as

\[
T(y; b, p) = \beta^{1/p} \left[ \sum_{i=1}^{K} a_i y_i^p \right]^{1/p},
\]

(12)

where \( a_i > 0, i = 1, \ldots, K, \) and

\[
\sum_{i=1}^{K} a_i = 1.
\]

Then

\[
\ln T = \frac{1}{p} \ln \beta + \frac{1}{p} \ln \left( \sum_{i=1}^{K} a_i y_i^p \right).
\]

(13)

a) First derivative

\[
\frac{\partial}{\partial p} \ln T = -\frac{1}{p^2} \ln \beta - \frac{1}{p^2} \ln \left( \sum_{i=1}^{K} a_i y_i^p \right) + \frac{1}{p} \cdot \frac{1}{\sum_{i=1}^{K} a_i y_i^p} \sum_{i=1}^{K} a_i y_i^p \ln y_i
\]

\[
= -\frac{1}{p^2} \sum_{i=1}^{K} a_i y_i^p \ln \left( \frac{y_i^p}{\beta \sum_{j=1}^{K} a_j y_j^p} \right), \quad p \neq 0.
\]

(14)

Since

\[
\frac{\partial}{\partial p} \ln T = \frac{1}{T} \frac{\partial T}{\partial p},
\]

vol. 29, n° 1, 1995
we immediately get

\[
\frac{\partial T}{\partial p} = \frac{\beta^{1/p}}{p^2} \left[ \sum_{i=1}^{K} a_i y_i^p \right]^{(1-p)/p} \left[ \sum_{i=1}^{K} a_i y_i^p \ln \left( \frac{y_i^p}{\beta \sum_{j=1}^{K} a_j y_j^p} \right) \right],
\]

\[p \neq 0.\]

It is interesting to note that for \( \beta > 1, \)

\[
\lim_{p \to 0^-} \frac{\partial T}{\partial p} = -\ln \beta \cdot \lim_{p \to 0^-} \left[ \frac{\beta^{1/p}}{p^2} \left( \sum_{i=1}^{K} a_i y_i^p \right)^{1/p} \right]
\]

\[= -\ln \beta \cdot \prod_{i=1}^{K} y_i^{a_i} \cdot \lim_{p \to 0^-} \left[ \frac{\beta^{1/p}}{p^2} \right] \quad \text{[equation (10)]}\]

\[= 0^-.
\]

Meanwhile, for \( \beta < 1, \) we obtain in similar fashion the following result.

\[
\lim_{p \to 0^+} \frac{\partial T}{\partial p} = 0^+.
\]

b) Second derivative

In the following summations \( i, j \in \{1, ..., K\} \) is understood, but omitted to simplify the notation.

\[
\frac{\partial^2}{\partial p^2} \ln T = \frac{2}{p^3} \ln \beta + \frac{2}{p^3} \ln \left( \sum_i a_i y_i^p \right) - \frac{2}{p^2} \sum_i a_i y_i^p \sum_i a_i y_i^p \ln y_i
\]

\[- \frac{1}{p} \cdot \left( \sum_i a_i y_i^p \right)^2 \left( \sum_i a_i y_i^p \ln y_i \right)^2
\]

\[+ \frac{1}{p} \sum_i a_i y_i^p \sum_i a_i y_i^p (\ln y_i)^2.
\]
After some re-arranging this reduces to

\[ \frac{\partial^2}{\partial p^2} \ln T = \frac{1}{p^3 \sigma^2} \left[ 2 \sigma \sum_i a_i y_i^p \ln \left( \frac{\beta \sigma}{y_i} \right) \right. \\
+ \left. p^2 \sum_{j<i} a_i a_j y_i^p y_j^p (\ln y_i - \ln y_j)^2 \right], \quad p \neq 0, \tag{18} \]

where

\[ \sigma = \sum_{i=1}^{K} a_i y_i^p. \tag{19} \]

3. PROPERTIES

Setting all the \( b_i = 1 \), we obtain the ordinary sum of order \( p \), denoted as follows:

\[ S(y; p) = \left( \sum_{i=1}^{K} y_i^p \right)^{1/p} \tag{20} \]

This sum satisfies the well-known relation,

\[ S(y; p_2) < S(y; p_1), \quad 0 < p_1 < p_2, \quad K \geq 2, \tag{21} \]

which is usually referred to as Jensen’s inequality. (For two different proofs of (21), see Theorem 19 in [6], and the proof by Beckenbach [1].) Beckenbach also shows that \( S(y; p) \) is convex in \( p \) for \( p > 0 \). His proof utilizes techniques from convex analysis. Using the results of the preceding section, we now obtain a general condition whereby Jensen’s inequality and the convexity result of Beckenbach [1] are extended to the weighted sum.

**Theorem 1**: Consider the function \( T(y; b, p) \) defined in (1), with given (constant) vectors \( y \) and \( b \). Assume without loss in generality that \( y_M = \max_i (y_i) \) occurs for a unique \( M \in \{1, ..., K\} \); i.e., there are no ties. (If this is not the case, add the coefficients \( (b_i) \) of the ties to form one term.) Then for \( K \geq 2 \), \( T(y; b, p) \) is a decreasing function of \( p \) for \( 0 < p < +\infty \), if, and only if, \( b_M \geq 1 \), where \( b_M \) is the coefficient of \( y_M \). Furthermore, if \( b_M \geq 1 \), \( T \) is also a strictly convex function in \( p \) over this interval.

vol. 29, n° 1, 1995
Proof:
(i) (If) Since \( p > 0 \), \( K \geq 2 \) and \( b_M > 1 \), it follows that

\[
\frac{y_i^p}{\beta \sigma} = \frac{y_i^p}{\sum_{j=1}^{K} b_j y_j^p} < \frac{y_i^p}{b_M y_M^p} \leq 1, \quad i = 1, \ldots, K. \tag{22}
\]

From equations (14) and (18), we see that

\[
\ln T \text{ is decreasing and strictly convex in } p \text{ for } 0 < p < +\infty. \text{ It immediately follows that } T \text{ is a decreasing function of } p \text{ in this interval. Furthermore,}
\]

\[
\frac{\partial^2 T}{\partial p^2} = T \frac{\partial^2}{\partial p^2} \ln T + \frac{1}{T} \left( \frac{\partial T}{\partial p} \right)^2; \tag{23}
\]

so that

\[
\frac{\partial^2 T}{\partial p^2} \geq T \frac{\partial^2}{\partial p^2} \ln T > 0.
\]

Thus, \( T \) is also strictly convex in \( p \), for \( 0 < p < +\infty \). We conclude that \( b_M \geq 1 \) is a sufficient condition for \( T \) to be a decreasing strictly convex function of \( p \in (0, +\infty) \).

(ii) (Only if) That \( b_M \geq 1 \) is a necessary condition for \( T \) to be decreasing in \( p \) immediately follows from the asymptotic behaviour of \( T \) as \( p \to +\infty \), shown in (5). If \( b_M < 1 \), \( T \) approaches \( y_M \) asymptotically from below, and hence is increasing and concave for sufficiently large \( p \).

Corollary 1: \( T(y; b, p) \) with \( K \geq 2 \) is a decreasing function of \( p > 0 \) for given weights \( b \) and all (positive) \( y \), if, and only if, \( b_i \geq 1 \), \( i = 1, \ldots, K \). Furthermore, \( T \) is also strictly convex in \( p \) under these conditions.

Proof: Consider any \( y \) such that the \( y_i \)'s are not all equal. Clearly, \( b_M \geq 1 \) if all the \( b_i \geq 1 \). By Theorem 1, we know that \( T \) is decreasing and strictly convex in \( p > 0 \). Now consider any \( y \) such that all the \( y_i \)'s are equal. Then, \( T = \beta^{1/p} y_1 \), where \( \beta = \sum_{i=1}^{K} b_i > 1 \) if all the \( b_i \geq 1 \). It is readily shown that \( T \) is once again decreasing and strictly convex in \( p > 0 \). Thus, \( b_i \geq 1 \),
$i = 1, \ldots, K$, is a sufficient condition. That this is also a necessary condition is readily seen by contradiction. Suppose $b_r < 1$, for some $r \in \{1, \ldots, K\}$. Construct a vector $y$ such that $y_i \neq y_j$, $i \neq j$, $\forall i, j \in \{1, \ldots, K\}$, and $y_r = \max_i (y_i)$. By the theorem, we know that $T$ is not a decreasing function of $p \in (0, +\infty)$, for this $y$.

The shape of $T$ as a function of $p$ becomes more complex when the criteria on the weights are changed, as shown in the following result.

**Property 1:** Consider a vector of weights $b$, such that $\beta = \Sigma_i b_i > 1$, and $b_r < 1$ for at least one $r \in \{1, \ldots, K\}$. Then, for any given $y$ there exists a $\delta > 0$ such that $T$ is decreasing and strictly convex in $p \in (0, \delta)$. However, if $y_r = \max_i (y_i)$, and there are no ties, then $T$ is increasing and strictly concave for sufficiently large positive $p$.

**Proof:** Follows immediately from the limit $p \rightarrow 0^+$ in (8) and the limit $p \rightarrow +\infty$ in (5).

Note that the function $T$ described in the preceding result is neither increasing or decreasing in $p$ nor convex or concave in $p$ over the entire interval $0 < p < +\infty$, and that at least one inflection point exists in this interval. This is illustrated in Figure 1.

The fact that $T$ is a decreasing function of $p$ in the interval $(0, +\infty)$ for all $y$ if, and only if, $b_i \geq 1$, $i = 1, \ldots, K$ (Corollary 1), has been recognized previously (Theorem 23, [6]). However, their proof is different than ours and does not show the important result that $T$ is convex in $p$ under these conditions. The third and final case to consider for the weights $b$ is where $\beta = \Sigma_i b_i \leq 1$. In the same theorem, the above authors prove that $T$ is non-decreasing in $p$ over the interval $(0, +\infty)$ for all $y$ if, and only if, this condition holds. Thus, the following property can be given without proof.

**Property 2:** A necessary and sufficient condition to have

$$T(y; b, p_1) \leq T(y; b, p_2), \quad 0 < p_1 < p_2,$$

for given weights $b$ and all $y$, is that $\beta \leq 1$. Furthermore, there is strict inequality unless all the $y_i$ are equal and $\beta = 1$. 

vol. 29, n° 1, 1995
An analogous result holds for negative values of \( p \). In this case, rewrite

\[
T(y; b, p) = \left[ \sum_{i=1}^{K} b_i y_i^p \right]^{1/p}
\]

\[
= 1 / \left[ \sum_{i=1}^{K} b_i \left( \frac{1}{y_i} \right)^{|p|} \right]^{1/|p|}
\]

\[ y = (1.5, 1, 1)^T \]
\[ b = (0.5, 1, 1)^T \]

![Figure 1. General shape of \( T \) under conditions of Property 1.](image)

It follows from Property 2 that

\[
T(y; b, p_1) \geq T(y; b, p_2), \quad p_2 < p_1 < 0,
\]

for given weights \( b \) and all \( y \), if and only if \( \beta \leq 1 \).

Use of negative \( p \) when the weighted sum \( T \) is a distance function in location models does not appear to have a physical interpretation. However, there may be other situations where \( p < 0 \) might be considered. In any case, we would like to take full advantage of our lengthy calculations of derivatives. This questionable motivation leads to the following results for \( p < 0 \).
THEOREM 2: Consider the function $T(y; b, p)$ defined in (1), with given (constant) vectors $y$ and $b$. Assume without loss in generality that $y_m = \min_i (y_i)$ occurs for a unique $m \in \{1, \ldots, K\}$; i.e., there are no ties. (If this is not the case, add the coefficients $(b_i)$ of the ties to form one term.) Then for $K > 2$, $T(y; b, p)$ is a decreasing function of $p$ in the interval $(-\infty, 0)$, if, and only if, $b_m \geq 1$, where $b_m$ is the coefficient of $y_m$. Furthermore, if $b_m \geq 1$, $\ln T$ is also a strictly concave function of $p$ over this interval.

Proof:

(i) (If). Since $p < 0$, $K \geq 2$, and $b_m \geq 1$, it follows that

$$\frac{y_i^p}{\sum_{j=1}^{K} b_j y_j^p} < \frac{y_m^p}{b_m y_m^p} = \frac{1}{b_m} \left( \frac{y_m}{y_i} \right)^{|p|} \leq 1, \quad i = 1, \ldots, K. \quad (24)$$

Returning to equations (14) and (18), we can readily show that

$$\frac{\partial \ln T}{\partial p} < 0 \quad \text{and} \quad \frac{\partial^2 \ln T}{\partial p^2} < 0, \quad p < 0.$$  

Hence, $\ln T$ is decreasing and strictly concave in $p$ in the interval $(-\infty, 0)$. It immediately follows that $T$ is a decreasing function of $p$ in this interval. We conclude that $b_m \geq 1$ is a sufficient condition for $T$ to be decreasing in $p$ and $\ln T$ strictly concave in $p$, for $-\infty < p < 0$.

(ii) (Only if). That $b_m \geq 1$ is a necessary condition for $T$ to be decreasing in $p$ immediately follows from the asymptotic behaviour of $T$ as $p \to -\infty$, shown in (6). If $b_m < 1$, $T$ approaches $y_m$ asymptotically from above; so that $T$ (or $\ln T$) is increasing and convex for sufficiently large negative values of $p$.

COROLLARY 2: $T(y; b, p)$ with $K \geq 2$ is a decreasing function of $p$ in the interval $(-\infty, 0)$ for given weights $b$ and all (positive) $y$, if, and only if, $b_i \geq 1$, $i = 1, \ldots, K$. Furthermore, $\ln T$ is strictly concave in $p$ under these conditions.

Proof: Consider any $y$ such that the $y_i$'s are not all equal. Clearly, $b_m \geq 1$ if all the $b_i \geq 1$. By Theorem 2, we know that $T$ is decreasing and $\ln T$ is strictly concave in $p \in (-\infty, 0)$. Now consider any $y$ such that all the $y_i$'s are equal. Then, $T = \beta^{1/p} y_1$, where $\beta = \Sigma_i b_i > 1$, if all the $b_i$'s $\geq 1$. It is readily shown once again that $T$ is decreasing and $\ln T$ is strictly
concave in \( p \in (-\infty, 0) \). We conclude that \( b_i \geq 1, i = 1, \ldots, K \), is a sufficient condition. That this is also a necessary condition is readily seen by contradiction, similar to the procedure in Corollary 1.

An open question posed by Beckenbach [1] concerns a lower bound on the number of inflection points of \( S(y; p) \) as a function of \( p \) in the interval \((-\infty, 0)\). This question is resolved below for the more general weighted sum.

**Theorem 3**: The function \( T(y; b, p) \) with \( \beta > 1 \) has at least one inflection point, and hence is neither convex nor concave in \( p \), in the interval \(-\infty < p < 0\).

*Proof*: From relations (6) and (8), it follows that \( T \) cannot be convex in \( p \) over the entire interval \((-\infty, 0)\). However, from (16) we see that a \( \delta > 0 \) exists such that \( T \) is convex in \( p \) in the interval \((-\delta, 0)\). Hence, we conclude that at least one inflection point exists.

It immediately follows that \( S(y; p) \) with \( K \geq 2 \) has at least one inflection point, and hence is neither convex nor concave in \( p \), in the interval \(-\infty < p < 0\). As a final comment we note that Theorem 3 applies when \( \beta < 1 \) and \( p \in (0, +\infty) \). The proof follows in a similar fashion from relations (5), (9) and (17).

4. CONCLUSIONS

This paper investigates properties of the weighted sum of order \( p \) as a function of the parameter \( p \). The weighted sum is a generalized form of the ordinary sum of order \( p \), which is alternatively known in location theory as the popular \( l_p \) distance function or \( l_p \) norm when \( p \geq 1 \).

The well-known result that the ordinary sum is a non-increasing, convex function of \( p \) in the interval \((0, +\infty)\) has been useful in deriving efficient algorithms to find the best-fitting parameter values of the weighted \( l_p \) norm for estimating distances in a given geographical region (*see* Brimberg and Love [3]). General conditions have been presented here for extending the monotonicity and convexity properties to the weighted sum. This sum can be interpreted as a generalized \( l_p \) distance which allows non-symmetric travel in the axial directions through the use of unequal weights on the coordinates. The new form should provide greater accuracy for estimating travel distances on a transportation network. However, there is a computational cost for the gain in accuracy, since additional parameters must be fitted to the network. For example, the weighted \( l_p \) norm in the plane has two unknown parameters,
whereas the generalized form discussed here will have three. The theoretical results presented in this paper are being applied by the authors to extend the existing fitting algorithms to the general weighted \( l_p \) norm. In addition, we plan to do empirical testing of the new metric.

This paper also resolves an open question in the literature concerning the existence of inflection points for sums of order \( p \) (see Beckenbach [1]). Since the conditions under which these inflection points occur are typically not encountered in practical applications of the sum of order \( p \) as a distance function, these particular results are currently of theoretical interest only.

ACKNOWLEDGEMENT

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and the Department of National Defense (Canada) Academic Research Program.

REFERENCES