

• UNIVERSITY.

MICHAEL G. DEGROOTE SCHOOL OF BUSINESS RESEARCH AND WORKING PAPER SERIES

ON THE DIRECTIONAL BIAS OF THE ℓ_{bp} -NORM By Halit Üster and Robert F. Love Michael G. DeGroote School of Business McMaster University Hamilton, Ontario Working Paper # 428 April, 1998 Innis Ontario, Canada L8S 4M4 (905) 525-9140 HB 74.5 .R47 no.428

ON THE DIRECTIONAL BIAS OF THE $\ell_{\scriptscriptstyle bp}$ -NORM

.

By

Halit Üster and Robert F. Love

Michael G. DeGroote School of Business McMaster University Hamilton, Ontario

Working Paper # 428

April, 1998

On the Directional Bias of the ℓ_{bp} -norm

Halit Üster

Robert F. Love

M.G. DeGroote School of Business Management Science/Systems Department McMaster University Hamilton, Ontario, L8S 4M4 Canada

Abstract

The Weighted Sum of Order p, the ℓ_{bp} -norm, is a generalization of the well-known ℓ_p -norm used in predicting distances in a transportation network. The properties of the directional bias function and the unit balls for the ℓ_{bp} -norm are of theoretical and practical interest. We investigate these properties and compare them with the properties of the ℓ_p -norm's directional bias function and the unit balls. We find that the ℓ_{bp} -norm is better at capturing the nonlinearity in a transportation network than the weighted ℓ_p -norm. It is also shown that, in contrast to the weighted ℓ_p -norm, where the optimal parameter p value is confined to the interval (1, 2), for the ℓ_{bp} -norm the parameter p can have an optimal value greater than 2.

Keywords: Distances, Directional Bias, Transportation.

1 Introduction

Distance predicting functions are involved in several different applications. In continuous location models, norms, as distance predicting functions, are usually employed in cost calculations to construct the objective function [13]. Such a model should represent the real situation as accurately as possible. Therefore the accuracy of the distance predicting function employed plays a crucial role in terms of the validity and the applicability of the model's output. Some other application areas which use distance predicting functions can be given as follows: distribution and transportation planning [5], [21]; accuracy validation of actual transportation network distance data [6]; response-time models for emergency vehicles [10]; construction of Voronoi Diagrams of a region [9]; location-allocation problems [13], and geographic information systems [17]. Moreover, the software packages Roadnet [15] and TruckStops2 [18] utilize distance predicting functions as a substitute to forming large files of distance data.

Love and Morris [11], [12] present several distance predicting functions which are mostly norms weighted by an inflation factor to account for the hills, bends and the other forms of "noise" in the transportation network. A significant conclusion of their study is that an empirical distance function should be tailored to a given region whenever a premium is placed on accuracy. This result is based on statistical analyses showing that the weighted ℓ_p -norm outperforms both the weighted Euclidean and the weighted rectangular norms.

In addition, it is shown by Love and Walker [14] that the weighted ℓ_p -norm is generally more accurate than a block norm [20]. The authors also observe that increasing the number of parameters of a block norm does not ensure that it becomes more accurate than the weighted ℓ_p -norm. Love and Morris [11] introduce the concept of axis rotation in their study on the road network in Milwaukee, Wisconsin. Brimberg, Love and Walker [4] investigate this concept in detail and conclude that a reference axis rotation chosen to align with the underlying pattern of the transportation network improves the accuracy of distance predictions. Huriot and Perreur [8] also discuss axis rotation and apply it in a study of the nine largest Swiss cities.

The functional form of the ℓ_{bp} -norm is given by the weighted sums of order p defined in Hardy et al. [7] (section 2.10). Üster and Love [19] show that the weighted sum of order p, $\ell_{bp}(\mathbf{x})$, is a norm and convex in \mathbf{x} . The use of the ℓ_{bp} -norm as a distance predicting function is first suggested by Brimberg and Love [3]. In the context of predicting travel distances we are mostly concerned with the characteristics of distance models in 2-dimensional Euclidean space. Therefore we define the ℓ_{bp} -norm as follows.

$$\ell_{bp}(\mathbf{x}) = (b_1 |x_1|^p + b_2 |x_2|^p)^{1/p}$$

where $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$. The parameters b_1, b_2 and p are generally assumed to be greater than zero. $\ell_{bp}(\mathbf{x})$ estimates the distance between any two points $\mathbf{y}, \mathbf{z} \in \mathbb{R}^2$ such that $\mathbf{x} = \mathbf{y} - \mathbf{z}$. The parameters b_1 and b_2 can be interpreted as non-symmetric weights along the axis directions in a distance or location model.

Brimberg and Love [3] state that, since it is a generalized form of the weighted ℓ_p -norm, the ℓ_{bp} -norm should provide greater accuracy for estimating distances on a transportation network. However, because of the extra parameter, we would expect an additional computational cost of fitting the ℓ_{bp} -norm to a transportation network.

Similar to the weighted ℓ_p -norm, the ℓ_{bp} -norm attempts to identify two inherent characteristics of a transportation network: rectangularity which is mostly associated with the parameter p, and nonlinearity which is associated with the parameters b_1 and b_2 . For example, a grid system or "Manhattan metric" type of road network has rectangularity properties and is captured by the parameter p ($p \approx 1$). On the other hand, a river or a mountain range or other irregularities will produce nonlinearity effects and will be explained by the parameters b_1 and b_2 [19].

In this paper we investigate some useful properties of the directional bias function of the ℓ_{bp} -norm. The way that the ℓ_{bp} -norm captures the rectangularity and the nonlinearity inherent in a transportation network is particularly important in the process of distance modelling and later in interpreting the parameters b_1 , b_2 and p. Using these properties we compare the ℓ_{bp} -norm and the well-known weighted ℓ_p -norm in terms of their ability to explain the underlying pattern in a transportation network and the procedures used to identify the best parameter values of each norm. Traditionally, the directional bias of norms is illustrated and compared by means of the unit balls associated with them.(see Fig.10.1 in [13]). For this reason we examine the unit ball of the ℓ_{bp} -norm as well as its directional bias function.

For the ℓ_p -norm, Brimberg and Love [2] conclude that any $\ell_q(\mathbf{x})$, q > 2, can be accurately approximated by a corresponding norm $\sigma \ell_p(\mathbf{x}')$, $1 , where <math>\sigma$ is a scaling factor and \mathbf{x}' gives the coordinates of \mathbf{x} after a 45° rotation of the reference axis. Therefore, for all practical purposes, the estimation of actual distances by an ℓ_p -norm with p > 2 need never be considered, since the same degree of accuracy can be obtained with a value of p in the interval (1, 2). In this study we find that for the ℓ_{bp} -norm, which is a generalization of the ℓ_p -norm, this fact does not generally hold. Whether an ℓ_{bq} -norm, q > 2, can be accurately approximated by another ℓ_{bp} -norm, where 1 , depends on the level of rectangularityand nonlinearity inherent in a transportation network. In the next section we define the directional bias function for the ℓ_{bp} -norm and present its properties. The last section is devoted to the implications of the results on distance modelling procedures.

2 **Properties**

Brimberg and Love [2] define directional bias for any norm k on \mathbb{R}^2 as

$$r(\theta) = \frac{k(\mathbf{x})}{\ell_2(\mathbf{x})}, \quad \mathbf{x} \neq 0, \quad \theta = \arctan\left(\frac{x_2}{x_1}\right).$$

Let u and v be two points in \mathbb{R}^2 such that $r(\theta_u) > r(\theta_v)$ and the same Euclidean distance is to be covered in both directions, i.e., $\ell_2(\mathbf{u}) = \ell_2(\mathbf{v})$. Then obviously $k(\mathbf{u}) > k(\mathbf{v})$ and we say that the difficulty of travel in the θ_u direction is greater than the difficulty of travel in the θ_v direction. Employing the ℓ_{bp} -norm as a distance predicting function we define the directional bias function as

$$r_{b_1,b_2,p}(\theta) = \frac{\ell_{bp}(\mathbf{x})}{\ell_2(\mathbf{x})} = \frac{(b_1|x_1|^p + b_2|x_2|^p)^{1/p}}{\ell_2(\mathbf{x})}$$
$$= (b_1|\cos\theta|^p + b_2|\sin\theta|^p)^{1/p}$$

where θ is the angle specifying the vector $\mathbf{x} \in \mathbb{R}^2$. We adopt the notation $r(\theta)$ to replace $r_{b_1,b_2,p}(\theta)$ in the rest of the paper.

Property 1 The directional bias is the same for two θ values 90° apart where the b_1 and b_2 values have been exchanged, p being the same.

Proof

$$r\left(\theta + \frac{\pi}{2}\right) = \left(b_1 \left| \cos\left(\theta + \frac{\pi}{2}\right) \right|^p + b_2 \left| \sin\left(\theta + \frac{\pi}{2}\right) \right|^p\right)^{1/p}$$

$$= (b_1|-\sin\theta|^p + b_2|\cos\theta|^p)^{1/p}$$
$$= (b_2|\cos\theta|^p + b_1|\sin\theta|^p)^{1/p}$$

so if we switch the b_1 and b_2 values, we obtain the same directional bias when θ is changed by $\pi/2$ with p being the same. \Box

Property 2 $r(\theta)$ is periodic with period π .

r

Proof

$$(\theta + \pi) = (b_1 | \cos(\theta + \pi)|^p + b_2 | \sin(\theta + \pi)|^p)^{1/p}$$
$$= (b_1 | -\cos\theta|^p + b_2 | \sin\theta|^p)^{1/p}$$
$$= (b_1 | \cos\theta|^p + b_2 | \sin\theta|^p)^{1/p}$$
$$= r(\theta)$$

and the result follows. $\hfill \Box$

Property 3 For any real w

$$r\left(\frac{\pi}{2} - w\right) = r\left(\frac{\pi}{2} + w\right)$$

i.e. $r(\theta)$ is the mirror image of itself about the line $\theta = \pi/2$.

Proof It follows from observing the equalities

$$\left|\sin\left(\frac{\pi}{2}-w\right)\right| = \left|\sin\left(\frac{\pi}{2}+w\right)\right|$$
 and $\left|\cos\left(\frac{\pi}{2}-w\right)\right| = \left|\cos\left(\frac{\pi}{2}+w\right)\right|$.

Thus we need to consider θ only in the interval $[0, \pi/2]$. \Box

In order to explore the shape of the $r(\theta)$ function we use the first- and the second-order derivatives. We next derive these derivatives. Note that $sin(\theta)$, $cos(\theta) \ge 0$ for $\theta \in [0, \pi/2]$.

$$\frac{dr(\theta)}{d\theta} = \frac{1}{(r(\theta))^{p-1}} (b_2 \cos\theta (\sin\theta)^{p-1} - b_1 \sin\theta (\cos\theta)^{p-1})$$
$$= \frac{\sin 2\theta}{2(r(\theta))^{p-1}} (b_2 (\sin\theta)^{p-2} - b_1 (\cos\theta)^{p-2}).$$

Observe that $dr(\theta)/d\theta = 0$ for $\theta = 0$ and $\theta = \pi/2$, and also for $\theta = \pi/4$ if $b_1 = b_2$.

$$\begin{aligned} \frac{d^2 r(\theta)}{d\theta^2} &= (r(\theta))^{1-2p} \frac{4}{(\sin 2\theta)^2} \left(-4b_1 b_2 \frac{(\sin 2\theta)^p}{2^p} + b_1 b_2 p(\sin \theta)^4 \frac{(\sin 2\theta)^p}{2^p} \right) \\ &- b_1^2 (\cos \theta)^{2p} \frac{(\sin 2\theta)^2}{4} + 2b_1 b_2 p \frac{(\sin 2\theta)^p}{2^p} \frac{(\sin 2\theta)^2}{4} \\ &- b_1 b_2 (\sin \theta)^4 \frac{(\sin 2\theta)^2}{2^p} - b_2^2 (\sin \theta)^{2p} \frac{(\sin 2\theta)^2}{4} \\ &+ b_1 b_2 p(\cos \theta)^4 \frac{(\sin 2\theta)^p}{2^p} - b_1 b_2 (\cos \theta)^4 \frac{(\sin 2\theta)^p}{2^p} \\ &= (r(\theta))^{1-2p} \left(\frac{(\sin 2\theta)^p}{2^{p-1}} b_1 b_2 (p-2) - b_1^2 (\cos \theta)^{2p} - b_2^2 (\sin \theta)^{2p} \\ &+ \left(\frac{(\sin 2\theta)}{2} \right)^{p-2} b_1 b_2 (p-1) ((\sin \theta)^4 + (\cos \theta)^4) \right) \end{aligned}$$

After some further rearrangements, we find that

$$\frac{d^2r(\theta)}{d\theta^2} = -r(\theta) + b_1b_2(p-1)(r(\theta))^{1-2p} \left(\frac{\sin 2\theta}{2}\right)^{p-2}$$

Property 4 $r(\theta)$ is differentiable and continuous for $\theta \in [0, \pi/2]$.

Proof We check the equality of the right- and the left-hand limits of $dr(\theta' + \epsilon)/d\theta$ as $\epsilon \to 0, \quad \theta' \in [0, \pi/2]$

$$\frac{dr(\theta)}{d\theta} = \frac{\sin 2(\theta' + \epsilon)(b_2 \sin^{p-2}(\theta' + \epsilon) - b_1 \cos^{p-2}(\theta' + \epsilon))}{2(b_1 \cos^p(\theta' + \epsilon) + b_2 \sin^p(\theta' + \epsilon))^{(p-1)/p}}.$$

Hence

$$\lim_{\epsilon \to 0^+} \left. \frac{dr(\theta)}{d\theta} \right|_{\theta = \theta' + \epsilon} = \frac{\sin 2\theta' (b_2 \sin^{p-2}\theta' - b_1 \cos^{p-2}\theta')}{2(r(\theta'))^{p-1}},$$

and

$$\lim_{\epsilon \to 0^-} \left. \frac{dr(\theta)}{d\theta} \right|_{\theta = \theta' + \epsilon} = \frac{\sin 2\theta' (b_2 \sin^{p-2}\theta' - b_1 \cos^{p-2}\theta')}{2(r(\theta'))^{p-1}}.$$

The equality of these limits shows that $r(\theta)$ is differentiable for $\theta \in [0, \pi/2]$. The continuity of $r(\theta)$ follows from the Theorems 4.7 and 4.9 in [16]. \Box

Next we give some properties of the $r(\theta)$ function by using the above derivatives. These properties are related to the shape of $r(\theta)$ for $\theta \in [0, \pi/2]$. The properties depend on the relative values of the parameters b_1 , b_2 and p. There are four possible cases relating these parameters:

- 1. $1 \le p < 2$, $b_1 < b_2$ (Property 5)
- 2. $1 \le p < 2$, $b_1 > b_2$ (Property 6)
- 3. p > 2, $b_1 < b_2$ (Property 7)
- 4. p > 2, $b_1 > b_2$ (Property 8)

For all four cases, the stationary point of $r(\theta)$, θ^* , is defined for $\theta \in [0, \pi/2]$ as follows:

$$\theta^* = \arctan\left(\frac{b_1}{b_2}\right)^{\frac{1}{p-2}}, \quad p \ge 1, \quad p \ne 2.$$

Notice that θ^* is a decreasing function of p when $b_1 > b_2$ and an increasing function of p when $b_1 < b_2$.

Property 5 If $1 \le p < 2$, and $b_1 < b_2$, then

a. $r(\theta)$ increases for $\theta \in (0, \theta^*)$, and decreases for $\theta \in (\theta^*, \pi/2)$, where $\theta^* \in [\pi/4, \pi/2]$.

b. $r(\theta)$ has two inflection points, $\tilde{\theta_1}$, $\tilde{\theta_2}$ where $\tilde{\theta_1} \in [0, \theta^*]$ and $\tilde{\theta_2} \in [\theta^*, \pi/2]$.

Proof

a. First notice that

$$\left. \frac{dr(\theta)}{d\theta} \right|_{\theta=0} = \left. \frac{dr(\theta)}{d\theta} \right|_{\theta=\frac{\pi}{2}} = 0.$$

For $\theta < \theta^*$ we have $\theta < \arctan(b_1/b_2)^{1/(p-2)}$ which implies that $b_2 (\sin\theta)^{p-2} - b_1 (\cos\theta)^{p-2} > 0$. Thus we obtain $dr(\theta)/d\theta > 0$, $\theta \in (0, \theta^*)$. Similarly $dr(\theta)/d\theta < 0$, for $\theta \in (\theta^*, \pi/2)$. Furthermore, since θ^* is the stationary point we have $dr(\theta^*)/d\theta = 0$.

Finally we observe that for $b_1 < b_2$ and $1 \le p < 2$, $tan\theta^* = (b_2/b_1)^{1/(2-p)} > 1$ which implies that $\theta^* \in [\pi/4, \pi/2]$.

b. First we consider the inflection point $\tilde{\theta_1}$. Since $1 \le p < 2$ we see that

$$\lim_{\theta \to 0^+} \frac{d^2 r(\theta)}{d\theta^2} = +\infty.$$

Furthermore, considering Property 4 and part (a) above it follows that

$$\left. \frac{d^2 r(\theta)}{d \theta^2} \right|_{\theta = \theta^*} < 0.$$

Hence we can state that $\exists \theta = \tilde{\theta_1} \ \ni \ d^2 r(\theta)/d \ \theta^2 = 0$, where $\theta \in [0, \theta^*], \ 1 \le p < 2$ and $b_1 < b_2$.

Next we consider the second inflection point $\tilde{\theta}_2$. It is readily known that

$$\left. \frac{d^2 r(\theta)}{d \theta^2} \right|_{\theta = \theta^*} < 0.$$

Furthermore, it can easily be verified that

$$\lim_{\theta \to \pi/2} \frac{d^2 r(\theta)}{d\theta^2} = +\infty.$$

Thus we conclude that $\exists \theta = \tilde{\theta_2} \quad \exists \quad d^2 r(\theta)/d\theta^2 = 0$, where $\theta \in [\theta^*, \pi/2], 1 \leq p < 2$ and $b_1 < b_2$. \Box

Property 6 If $1 \le p < 2$, and $b_1 > b_2$, then

a. $r(\theta)$ increases for $\theta \in (0, \theta^*)$, and decreases for $\theta \in (\theta^*, \pi/2)$, where $\theta^* \in [0, \pi/4]$.

b. $r(\theta)$ has two inflection points, $\tilde{\theta_1}$, $\tilde{\theta_2}$ where $\tilde{\theta_1} \in [0, \theta^*]$ and $\tilde{\theta_2} \in [\theta^*, \pi/2]$.

Proof

a. Similar to Property 5.a we have $dr(\theta)/d\theta > 0$, where $\theta \in (0, \theta^*)$, $dr(\theta)/d\theta < 0$, for $\theta \in (\theta^*, \pi/2)$, and $dr(\theta^*)/d\theta = 0$. However, in this case since $b_1 > b_2$, we have $tan\theta^* = (b_2/b_1)^{1/(2-p)} < 1$ which implies that $\theta^* \in [0, \pi/4]$.

b. We can use the same approach that was used in Property 5 to examine the inflection points because we still have the case where $1 \le p < 2$. On the other hand, we can make use of Properties 1 and 3. In this case it is not necessary to use the derivatives, but instead we utilize the relations between functions $r(\theta)$ with $1 \le p < 2$, $b_1 < b_2$, and $r(\theta)$ with $1 \le p < 2$, $b_1 > b_2$. \Box

Property 7 If p > 2, and $b_1 < b_2$, then

a. $r(\theta)$ decreases for $\theta \in (0, \theta^*)$, and increases for $\theta \in (\theta^*, \pi/2)$, where $\theta^* \in [0, \pi/4]$.

b. $r(\theta)$ has two inflection points, $\tilde{\theta_1}$, $\tilde{\theta_2}$ where $\tilde{\theta_1} \in [0, \theta^*]$ and $\tilde{\theta_2} \in [\theta^*, \pi/2]$.

Proof

a. We again make use of the first derivative of $r(\theta)$. Notice that

$$\frac{dr(\theta)}{d\theta}\Big|_{\theta=0} = \frac{dr(\theta)}{d\theta}\Big|_{\theta=\frac{\pi}{2}} = 0.$$

For $\theta < \theta^*$ we have $\theta < \arctan(b_1/b_2)^{1/(p-2)}$ which implies that $b_2 (\sin\theta)^{p-2} - b_1 (\cos\theta)^{p-2} > 0$. Thus we obtain $dr(\theta)/d\theta > 0$, $\theta \in (0, \theta^*)$. Similarly for $\theta \in (\theta^*, \pi/2)$, $dr(\theta)/d\theta < 0$. Also note that since θ^* is the stationary point we have $dr(\theta^*)/d\theta = 0$.

Finally we observe that for $b_1 < b_2$ and p > 2, $tan\theta^* = (b_1/b_2)^{1/(p-2)} > 1$ implying that $\theta^* \in [0, \pi/4]$.

b. First consider the inflection point $\tilde{\theta_1}$. It can easily be verified that

$$\left. \frac{d^2 r(\theta)}{d \theta^2} \right|_{\theta=0} = -b_1^{1/p} < 0.$$

Furthermore, it follows from Property 4 and the first part of this property that

$$\left. \frac{d^2 r(\theta)}{d \theta^2} \right|_{\theta = \theta^*} > 0.$$

Thus we conclude that $\exists \theta = \tilde{\theta_1} \ \ni \ d^2r(\theta)/d\,\theta^2 = 0$, where $\theta \in [0, \theta^*], \ p > 2$ and $b_1 < b_2$.

Next we consider the second inflection point $\tilde{\theta}_2$ which is in the interval $[\theta^*, \pi/2]$. It is already known that

$$\left. \frac{d^2 r(\theta)}{d \theta^2} \right|_{\theta = \theta^*} > 0.$$

Moreover, it can be shown that

$$\left. \frac{d^2 r(\theta)}{d \theta^2} \right|_{\theta=\pi/2} = -b_2^{1/p},$$

and therefore the second derivative of $r(\theta)$ at $\theta = \pi/2$ is negative. Thus we conclude that $\exists \theta = \tilde{\theta_2} \ \ni \ d^2 r(\theta)/d \theta^2 = 0$, where $\theta \in [\theta^*, \pi/2], p > 2$ and $b_1 < b_2$. \Box

Property 8 If p > 2, and $b_1 > b_2$, then

a. $r(\theta)$ decreases for $\theta \in (0, \theta^*)$, and increases for $\theta \in (\theta^*, \pi/2)$, where $\theta^* \in [\pi/4, \pi/2]$.

b. $r(\theta)$ has two inflection points, $\tilde{\theta_1}$, $\tilde{\theta_2}$ where $\tilde{\theta_1} \in [0, \theta^*]$ and $\tilde{\theta_2} \in [\theta^*, \pi/2]$.

Proof

a. As in Property 7.a we have $dr(\theta)/d\theta < 0$, where $\theta \in (0, \theta^*)$, $dr(\theta)/d\theta > 0$, for

 $\theta \in (\theta^*, \pi/2)$, and $dr(\theta^*)/d\theta = 0$. However, in this case where $b_1 > b_2$, we have $tan\theta^* = (b_1/b_2)^{1/(p-2)} > 1$ showing that $\theta^* \in [\pi/4, \pi/2]$.

b. Similar to Property 6.b, we can use two approaches to examine the inflection points. We can employ the same approach that we have used for Property 7 since we still have the case where p > 2. Secondly, we can make use of Property 1 and Property 3. In this case instead of using the derivatives we utilize the relations between functions $r(\theta)$ with p > 2, $b_1 < b_2$ and $r(\theta)$ with p > 2, $b_1 > b_2$. \Box

We have already mentioned that depending on the relative values of the parameters b_1 , b_2 , and p there are four possible shapes of the $r(\theta)$ function. In Properties 5 to 8 we have identified these shapes.

Observe that the boundary values of $r(\theta)$ are

$$r(0) = b_1^{1/p}$$
 and $r(\pi/2) = b_2^{1/p}$,

and $r(\theta)$ evaluated at the stationary point θ^* is given by

$$r(\theta^*) = \left[b_1 \left(1 + (b_1/b_2)^{2/(p-2)} \right)^{-1/2p} + b_2 \left(\frac{(b_1/b_2)^{1/(p-2)}}{\left(1 + (b_1/b_2)^{2/(p-2)} \right)^{1/2}} \right)^p \right]^{1/p}$$

These expressions directly follow from substituting the values of θ ; θ^* , 0, and $\pi/2$, in $r(\theta)$.

We give the following property without an explicit proof. It follows from the definition of $r(\theta)$ and Corollary 1 in Brimberg and Love [3].

Property 9 Let b_1 and b_2 be given parameter values and p > 1. Then $r(\theta)$ is a decreasing function of p for any fixed $\theta \in [0, \pi/2]$.

Note that, in contrast to the directional bias function of the ℓ_p -norm, this property is also valid at the boundaries, i.e., $\theta = 0$ and $\theta = \pi/2$. As a direct consequence of this result we state the following property.

Property 10 Consider two $r(\theta)$ functions, $r_1(\theta)$, $r_2(\theta)$, with a set of given b_1 and b_2 values and the parameter p, such that $1 \le p < 2$ for $r_1(\theta)$ and p > 2 for $r_2(\theta)$. Then we have

$$r_1(\theta) \ge r_2(\theta), \quad \theta \in \left(0, \frac{\pi}{2}\right).$$

Observe that the equality above holds only if $b_1 = 1$ in which case $r_1(0) = r_2(0) = 1$.

We note the following limiting cases related to the stationary point $(\theta^*, r(\theta^*))$ which will be useful in the subsequent discussion. Particularly, we are interested in the behaviour of θ^* and $r(\theta^*)$ when $p \to 2$ and also $b_1 \to b$, $b_2 \to b$ or $(b_1/b_2) \to 1$. We present the limiting cases in four groups along with the conditions examined in Properties 5 to 8.

1.
$$1 \le p < 2$$
, $b_1 < b_2$: (Property 5)
$$\lim_{p \to 2^-} \theta^* = \frac{\pi}{2}, \quad \lim_{p \to 1} \theta^* = \arctan\left(\frac{b}{b_1}\right), \quad \lim_{\frac{b_1}{b_2} \to 1^-} \theta^* = \frac{\pi}{4}, \quad \lim_{p \to 2^-} r(\theta^*) = b_2^{1/p}.$$

2. $1 \le p < 2$, $b_1 > b_2$: (Property6)

$$\lim_{p \to 2^{-}} \theta^{*} = 0, \quad \lim_{p \to 1} \theta^{*} = \arctan\left(\frac{b}{b_{1}}\right), \quad \lim_{\frac{b_{1}}{b_{2}} \to 1^{+}} \theta^{*} = \frac{\pi}{4}, \quad \lim_{p \to 2^{-}} r(\theta^{*}) = b_{1}^{1/p}.$$

3. p > 2, $b_1 < b_2$: (Property 7)

$$\lim_{p \to 2^+} \theta^* = 0, \quad \lim_{p \to +\infty} \theta^* = \frac{\pi}{4}, \quad \lim_{\frac{b_1}{b_2} \to 1^-} \theta^* = \frac{\pi}{4}, \quad \lim_{p \to 2^+} r(\theta^*) = b_1^{1/p}.$$

4. p > 2, $b_1 > b_2$: (Property 8)

$$\lim_{p \to 2^+} \theta^* = \frac{\pi}{2}, \quad \lim_{p \to +\infty} \theta^* = \frac{\pi}{4}, \quad \lim_{\frac{b_1}{b_2} \to 1^+} \theta^* = \frac{\pi}{4}, \quad \lim_{p \to 2^+} r(\theta^*) = b_2^{1/p}.$$

Finally, valid for all cases we observe that

$$\lim_{b_1 \to b} \lim_{b_2 \to b} r(\theta^*) = b^{\frac{1}{p}} 2^{\frac{1}{p} - \frac{1}{2}}.$$

The plots of $r(\theta)$ with different parameter values b_1 , b_2 and p are given in Figure 1. It can easily be seen from the plots that the ℓ_{bp} -norm models directional bias in a quite different way than the ℓ_p -norm (see Figure 1 in [2]).

The unit ball of the ℓ_{bp} -norm is given in Figure 2. For clarity, we show two sets of unit ball plots corresponding to Properties 5 to 8. One immediate observation is that, in contrast to the unit ball of the ℓ_p -norm (see [12]), the x_1 - and x_2 -axis intercepts are not equal to 1, but are given by $b_1^{-1/p}$ and $b_2^{-1/p}$, respectively. Consider the case where $b_1 \neq b_2$ and p = 2. Then the unit ball actually becomes an ellipse with the equation

$$(b_1 |x_1|^2 + b_2 |x_2|^2)^{1/2} = 1.$$

For values of p decreasing from two to one this elliptic shape of the unit ball shrinks and ultimately becomes a diamond-shaped ball for p = 1. Conversely for values of p increasing from two to infinity the unit ball expands and converges to a unit square. That is, it becomes symmetric with respect to orthogonal coordinate axes. In a sense we can say that a relatively high degree of rectangularity offsets the directional nonlinearity captured by the parameters b_1 and b_2 . Note also that since b_1 , $b_2 > 1$, all unit balls for p > 1 will be enclosed in this unit square.

3 Implications for Distance Prediction

The directional bias function, $r(\theta)$, provides some practical insights which are useful in modelling distances. Distance modelling in a region basically involves determining the best



a. $b_1 = 1.1, b_2 = 1.6, A: p = 1.4, B: p = 1.7, C: p = 1.9, D: p = 2.0, E: p = 2.1, F: p = 2.6, G: p = 4.0$



b. $b_1 = 1.2, b_2 = 1.1, A : p = 1.2, B : p = 1.6, C : p = 1.8, D : p = 2.0, E : p = 2.2, F : p = 2.8, G : p = 4.5$

Figure 1: $r(\theta)$ plots



a. $b_1 = 1.1, b_2 = 1.6, A : p = 1.0, B : p = 1.5, C : p = 2.0, D : p = 2.5, E : p = 4$

~



b. $b_1 = 2.6, b_2 = 1.1, A : p = 1.0, B : p = 1.5, C : p = 2.0, D : p = 2.5, E : p = 4$ Figure 2: ℓ_{bp} -norm Unit Balls

parameter values of the distance predicting function, e.g. b_1 , b_2 and p in the ℓ_{bp} -norm, so that a prediction-errors-related goodness-of-fit criterion value is minimized. Computational procedures to determine the best parameter values are given for the weighted ℓ_p -norm by Brimberg and Love [1] and for the ℓ_{bp} -norm by Üster and Love [19]. In both procedures, for the parameter p it is necessary to conduct a search over a safe range of values that includes the optimal value of p for the transportation network. Brimberg and Love [2] analyze the directional bias function of the ℓ_p -norm, $r_p(\theta)$, in detail and conclude that for all practical purposes the estimation of actual distances by an ℓ_p -norm with p > 2 need never be considered, since the same degree of accuracy can be obtained with a value of p in the interval [1, 2] after rotating the axes by 45°.

We define the direction of greatest (least) difficulty θ_g (θ_l) as the value of θ which maximizes (minimizes) the $r(\theta)$ function. Let x and y be two points in R^2 separated by a straight line segment \mathcal{L} of fixed length $\ell_2(x-y)$. Then $\ell_{bp}(x-y)$ is maximized if \mathcal{L} is parallel to θ_g and minimized if \mathcal{L} is parallel to θ_l . It is shown by Brimberg and Love [2] that for the ℓ_p -norm, which is a special form of the ℓ_{bp} -norm with $b_1 = b_2 = 1$, $\theta_g = \pi/4$, $\theta_l = 0$, $\pi/2$ for $1 \leq p < 2$ and $\theta_g = 0$, $\pi/2$, $\theta_l = \pi/4$ for p > 2. For the ℓ_{bp} -norm, inspecting the graph of $r(\theta)$ (Figure 1) and the unit ball (Figure 2), we see that when $1 \leq p < 2$, $\theta_l = 0$ for $b_1 < b_2$ and $\theta_l = \pi/2$ for $b_1 > b_2$, whereas the direction of greatest difficulty θ_g is such that $\theta_g \in (\arctan(b_2/b_1), \pi/2)$ for $b_1 < b_2$ and $\theta_g \in (0, \arctan(b_2/b_1))$ for $b_1 > b_2$. For p > 2 the situation is somewhat similar but in the opposite sense. In particular, the direction of least difficulty θ_l is such that $\theta_l \in (0, \pi/4)$ for $b_1 < b_2$ and $\theta_l \in (\pi/4, \pi/2)$ for $b_1 > b_2$, while the direction of greatest difficulty is at $\theta_g = \pi/2$ for $b_1 < b_2$ and $\theta_g = 0$ for $b_1 > b_2$. As a result of this non-fixed θ_g for $1 \leq p < 2$ and θ_l for p > 2, a phase change in the directions of greatest

ť

and least difficulty does not occur in the strict sense that it occurs for the ℓ_p -norm at $\theta = \pi/4$. In other words, we are not in general guaranteed that θ_g for $1 \le p < 2$ and θ_l for p > 2 coincide as in the ℓ_p -norm where they are both equal to $\pi/4$.

Suppose that we have a transportation network where an ℓ_{bp} -norm with $1 \leq p < 2$ can be closely approximated by an ℓ_{bq} -norm where q > 2 and the coordinate axes are rotated. Then while modelling distances in this network we have two minimum criterion values for $\theta \in [0, \pi/2]$, one with $1 \leq p < 2$ and the other with p > 2. We will call these minimum criterion values 'bottoms' in the *Criterion vs.* θ graphs. These bottoms correspond to SD_1 and SD_2 in Table 1 and Figure 4 for the example given at the end of this section. In that example we use the 'Sum of Squared Deviations' (*SD*) as the criterion [11]. If the weighted ℓ_p -norm is used, then the above mentioned close approximation is always possible and therefore the bottoms always occur 45° apart with the same minimum criterion values. This is a direct consequence of the exact phase change of 45° observed in the $r_p(\theta)$ graph (see Figure 1 in [2]). However for the ℓ_{bp} -norm, this approximation does not necessarily exist for a given transportation network.

Before proceeding to the discussion on this difference between the weighted l_p -norm and the l_{bp} -norm, we define an *indicator of directional nonlinearity*, τ , at an axis rotation as follows:

$$\tau = \frac{max\{b_1, b_2\}}{min\{b_1, b_2\}},$$

where b_1 and b_2 are the best parameter values for an axis rotation θ . The corresponding τ values for SD_1 and SD_2 are denoted by τ_1 and τ_2 , respectively. We can employ $\Delta \tau = |\tau_1 - \tau_2|$ as an indicator for the existence of directional nonlinearity in the transportation network. While a high value of $\Delta \tau$ indicates the existence of directional nonlinearity, a low value presents





evidence for uniform nonlinearity. For example, consider a perfectly rectangular transportation network, aligned with the conventional coordinate axis, with an irregularity, say a mountain range, in the vertical direction as shown in Figure 3. It is clear that both nonlinearity and rectangularity are highly pronounced characteristics of the underlying pattern. When fitting distances in such a transportation network we expect to see a bottom in the *SD* vs. θ graph at $\theta = 0$ with the parameter p value very close to 1. b_1 will be relatively high compared to b_2 showing the high level of nonlinearity or difficulty of travel in x_1 direction. τ_1 will be high indicating the directional nonlinearity captured at this axis rotation. The second bottom occurs after 45° of axis rotation and corresponds to a p value much greater than 2. This large value of p still captures the rectangularity accurately in the network [2]. Since the irregularity has the same effect on travel in both directions after the axis rotation, the b_1 and b_2 values will be very close. Relatively high values of both b_1 and b_2 will still indicate the high level of nonlinearity. However, the low value of τ_2 , close to 1, will show an existence of uniform nonlinearity at this axis rotation. The resulting high values of $\Delta \tau$ will indicate the existence of a predominant direction of nonlinearity in this particular transportation network. When $\theta = 0$, the axes are in perfect alignment with both rectangularity and directional nonlinearity inherent in the network, and the level of rectangularity, nonlinearity and the existence of directional nonlinearity are modelled accurately. After the axis rotation we will have the worst alignment with the underlying rectangularity and directional nonlinearity. When $\theta = 45^{\circ}$, the distance model still captures the level of rectangularity and nonlinearity. However, τ_2 does not indicate the existing directional nonlinearity in the transportation network. As a result, we can not expect to obtain the same SD_1 and SD_2 values for such a network; SD_1 will be much lower than SD_2 . The road patterns in Sydney (Table 1), and Pennsylvania, London North and Los Angeles [19], are examples for this case. Conversely, if we have uniform nonlinearity over all the transportation network, then τ will be insensitive to axis rotation so that $\Delta \tau$ will be very close to zero as in the case of Toronto (Table 1).

Returning to our discussion on the existence of two bottoms in the SD vs. θ graph, we first consider only the boundaries $\theta = 0, \pi/2$. It follows from Figure 1 and Property 1 that if the criterion (SD) attains its minimum at an axes rotation θ , then the same minimum criterion value must be obtained at $(\theta + \pi/2)$ where b_1 and b_2 values are exchanged, p being the same. We call this minimum value of the criterion occurring in the *Criterion vs.* θ graph the first bottom. Secondly we consider the cases in which another minimum criterion value, the second bottom, is attained in the interval $(0, \pi/2)$. We argue that the existence of such a second bottom in the *Criterion vs.* θ graph depends on the existence of directional nonlinearity and the level of rectangularity in the transportation network. If the underlying pattern is highly Euclidean, i.e. $p \approx 2$, then all four limiting cases suggest that a possible second bottom is not likely to occur. In this case the θ_i 's approach the boundaries making the phase change of $\pi/2$ discussed above more likely to occur. The best parameter p value can then occur either in the interval (1, 2)or $(2, +\infty)$. We can say that in such a case the directional nonlinearity clearly dominates the rectangularity in the network. On the other hand, if there exists a less pronounced directional nonlinearity, i.e. $\Delta \tau \approx 0$, or a high rectangularity, i.e. $p \to 1$ or $p \to +\infty$, then θ_g for $1 \le p < 2$ and θ_l for p > 2 move towards $\pi/4$ or equivalently the unit ball actually converges to a shape similar to the unit ball of the weighted ℓ_p -norm. Hence this case actually resembles the weighted ℓ_p -norm and a second bottom is likely to occur in the *Criterion vs.* θ graph. This time the rectangularity dominates the directional nonlinearity in the network. The second bottom which occurs as a result of approximately a 45° phase change on the $r(\theta)$ graph is about 45° apart from the first bottom.

These observations lead us to the conclusion that while modelling distances by using the ℓ_{bp} -norm we may not always be assured that there exists a good fit with a parameter pvalue in the interval (1, 2). Therefore we can not limit our search for the optimal p value of the norm to this interval. However, we can still impose an upper bound on the search range of p. Suppose that the underlying pattern has a dominant rectangularity. Then we know that there exists two sets of best parameter values where one of them has a p value in the interval (1, 2). Therefore we have an upper bound of 2 for the search range of p in this case. Now suppose that there exists a predominant direction of nonlinearity in the transportation network. Because of the possibility of having an optimal parameter p value greater than 2 for such a case, we have to consider a larger search range, $(1, \bar{p})$ where $\bar{p} > 2$, in distance modelling algorithms. This upper bound must be high enough to capture the rectangularity in the transportation network so that if the optimal p is above this limit there will be a corresponding optimal fit with a pvalue close to 1. Although the choice of this upper bound for the search range of p is left to the

ť

	SD_1 (First Bottom)	SD_2 (Second Bottom)	$\Delta S D$
REGION	θ	θ	
	p, b_1, b_2	p, b_1, b_2	
	$ au_1$	$ au_2$	Δau
	172.34		
Great Britain	0°	-	-
	2.0352, 1.1116, 1.3925		
	1.2527		
	1.0984	1.3550	23.26%
Sydney	8°	54°	
	1.5571, 1.3675, 1.1521	3.0007, 2.1130, 2.1541	
	1.1870	1.0195	0.1675
	5.0557	5.0619	0.12%
Toronto	42°	87°	
1	5.2215, 4.1261, 4.2427	1.2009, 1.0581, 1.0275	
	1.0282	1.0298	0.0016

Table 1: Example Parameter Values

analyst's preliminary inspection of the transportation network analysed, our empirical work on seventeen geographic regions [19] reveals that use of a search range [1, 4] will always obtain the best parameter p value.

As an example, we comment on the optimum parameter values for three of the above mentioned regions, Great Britain, Sydney and Toronto. We are mainly interested in the interpretation of the optimal parameter values which are given in Table 1. The *SD vs.* θ graphs are given in Figure 4. Great Britain is a case where the optimal p is greater than 2. Since p is very close to 2 the road network is highly Euclidean. The value of τ_1 (1.2527) represents the high directional nonlinearity modelled at the axis rotation $\theta = 0$. There exists only one minimum criterion value occurring in the $[0, \pi/2]$ interval. In Sydney's case we observe two bottoms, and there is a considerable gap between them, 23.36%. This is because the region is fairly rectangular (p = 1.5571) but at the same time there exists directional nonlinearity. However, neither is dominant. The directional nonlinearity in the transportation network is evident from the significant value of $\Delta \tau$ (0.1675). The first bottom better represents the existing



ť

Ì

Figure 4: *SD vs.* θ Plots

directional nonlinearity with its higher τ_1 value and thus provides a lower *SD* criterion value than the second bottom. For Toronto, there exists a low level of directional nonlinearity with $\tau_1 = 1.0282$ and $\Delta \tau = 0.0016$, and the network is highly rectangular. The rectangularity in the road pattern clearly dominates the directional nonlinearity. Hence, we see the two bottoms 45° apart and there is a negligible amount of gap (0.12%) between the criterion values at the bottoms.

....

1

Acknowledgements : This research was supported by a grant and a scholarship from the Natural Sciences and Engineering Research Council of Canada.

References

- Brimberg, J., and Love, R.F., "Estimating Travel Distances by the Weighted l_pnorm", Naval Research Logistics, 38, (1991), pp. 241-259.
- Brimberg, J. and Love, R.F., "Directional Bias of the lp-norm", European Journal of Operational Research, 67, (1993), pp. 287-294.
- Brimberg, J., Love, R.F., "Properties of Ordinary and Weighted Sums of Order p
 Used for Distance Estimation", Recherche Opérationnelle, 29(1), (1995), pp. 59-72.

f

- [4] Brimberg, J., Love, R.F., Walker, J.H., "The effect of axis rotation on distance estimation", European Journal of Operational Research, 80, (1995), pp. 357-364.
- [5] Eilon, S., Watson-Gandy, C.D.T., and Christofides, N., Distribution Management: Mathematical Modeling and Practical Analysis, Hafner Publishing Company, New York (1971).
- [6] Ginsburgh, V. and Hansen, P., "Procedures for the Reduction of Errors in Road Network Data", Operational Research Quarterly, 25, (1974), pp. 321-322.
- [7] Hardy, G.H., Littlewood, J.E., Pólya, G., *Inequalities*, 2nd Edition, Cambridge University Press, Cambridge (1952).

- [8] Huriot, J.M. and Perreur, J., "On the Weber Problem with Rectangular Distance: A Comment", Management Science, 20, (1973), pp. 418-419.
- [9] Klein, R., "Voronoi Diagrams in the Moscow Metric", Proc. Graph-Theoretic Concepts in Computer Science (WG '88), Amsterdam, (1988), pp. 434-441.
- [10] Kolesar, P., Walker, W., and Hausner, J., "Determining the Relation between Fire Engine Travel Times and Travel Distances in New York City", Operations Research, 23, (1975), pp. 614-627.
- [11] Love, R.F., and Morris, J.G., "Modeling Inter-City Road Distances by Mathematical Models", Operational Research Quarterly, 23, (1972), pp. 61-71.
- [12] Love, R.F., and Morris, J.G., "Mathematical Models of Road Travel Distances", Management Science, 25, (1979), pp. 130–139.
- [13] Love, R.F., Morris, J.G., and Wesolowsky, G.O., Facilities Location Models and Methods, North-Holland, New York (1988).
- [14] Love, R.F., Walker, J.H., "An Empirical Comparison of Block and Round Norms for Modeling Actual Distances", *Location Science*, 2, No.1, (1994), pp. 21–43.
- [15] ROADNET, Roadnet Technologies, 10540 York Rd., Huntvalley, Maryland, 1993.
- [16] Rudin, W., Principles of Mathematical Analysis, McGraw-Hill, Inc., New York, (1976).
- [17] Star, J. and Estes, J., Geographic Information Systems An Introduction, Prentice-Hall, New Jersey (1990).

- [18] TruckStops 2, MicroAnalytics, Suite One, 2045 North 15th Street, Arlington, Virginia and 1986 Queen Street East, Toronto, Ontario (1993).
- [19] Üster, H., Love, R.F., "Application of Weighted Sums of Order p to Distance Estimation", Faculty of Business Working Paper, McMaster University, Hamilton, Ontario, Canada (1998).
- [20] Ward, J.E., and Wendell, R.E., "Using Block Norms for Location Modeling", Operations Research, 33, (1985), pp. 1074–1090.
- [21] Westwood, J.B., "A Transport Planning Model for Primary Distribution", *Interfaces*, 8, (1977), pp.1–10.

ŧ

Faculty of Business McMaster University

WORKING PAPERS - RECENT RELEASES

- 401. Ali R. Montazemi and Feng Wang, "An Empirical Investigation of CAI in Support of Mastery Learning", February, 1995.
- 402. Kalyan Moy Gupta and Ali Reza Montazemi, "Retrieval in Case-Based Reasoning Systems with Modified Cosine Matching Function", February, 1995.
- 403. Kalyan Moy Gupta and Ali Reza Montazemi, "A Connectionist Approach for Similarity Assessment in Case-Based Reasoning Systems", March, 1995.
- 404. John W. Medcof, "Selecting Alliance and Network Partners Strategic Issues", March 1995.
- 405. Jiang Chen and George Steiner, "Discrete Lot Streaming in Two-Machine Flow Shops", March, 1995.
- 406. Harish C. Jain and S. Muthuchidambaram, "Strike Replacement Ban in Ontario and Its Relevance to U.S. Labor Law Reform", May, 1995.
- 407. Ali R. Montazemi and Kalyan Moy Gupta, "A Framework for Retrieval in Case-Based Reasoning Systems", June, 1995.
- 408. Ali R. Montazemi and Kalyan Moy Gupta, "An Adaptive Agent for Case Description in Diagnostic CBR Systems", June, 1995.
- 409. Roy J. Adams, Noel Cowell and Gangaram Singh, "The Malaing of Industrial Relations in the Commonwealth Caribbean", June, 1995.
- 410. Jiang Chen and George Steiner, "Approximation Methods for Discrete Lot Streamingin Flow Shops", June, 1995.
- 411. Harish C. Jain and S. Muthuchidambaram, "Bill 40 Amendments to Ontario Labour Relations Act: An Overview and Evaluation", June, 1995.
- 412. Jiang Chan and George Steiner, "Discrete Lot Streaming in Three-Machine Flow Shops", July, 1995.
- 413. J. Brimberg, A. Mehrez and G.O. Wesolowsky, "Allocation of Queueing Facilities Using a Minimax Criterion", January, 1996.

- 414. Isik Zeytinoglu and Jeanne Norris, "Global Diversity in Employment Relationships: A Typology of Flexible Employment", March, 1996.
- 415. N. Archer, "Characterizing World Wide Web Search Strategies", April, 1996.
- 416. J. Rose, "Immediacy and Saliency in Remedying Employer Opposition to Union Organizing Campaigns", July, 1996.
- 417. Roy J. Adams and Parbudyal Singh, "Worker Rights Under NAFTA: Experience With the North American Agreement on Labor Cooperation", September, 1996.
- 418. George Steiner and Paul Stephenson, "Subset-Restricted Interchange for Dynamic Min-Max Scheduling Problems", September, 1996.
- 419. Robert F. Love and Halit Uster, "Comparison of the Properties and the Performance of the Criteria Used to Evaluate the Accuracy of Distance Predicting Functions", November, 1996.
- 420. Harish C. Jain and Simon Taggar, "The Status of Employment Equity in Canada", December, 1996.
- 421. Harish C. Jain and Parbudyal Singh, "Beyond The Rhetoric: An Assessment of the Political Arguments and Legal Principles on Strike Replacement Laws in North America", January, 1997.

4

8

5

- 422. Jason Schwandt, "Electronic Data Interchange: An Overview of Its Origins, Status, and Future", March, 1997.
- 423. Christopher K. Bart with John C. Tabone, "Mission Statement Rationales and Organizational Alignment in the Not-for-profit Healthcare Sector", November, 1997.
- 424. Harish C. Jain, Michael Piczak, Işik Urla Zeytinoğlu, "Workplace Substance Testing An Exploratory Study", November, 1997.
- 425. S. Suarga, Yufei Yuan, Joseph B. Rose, and Norman Archer, "Web-based Collective Bargaining Support System: A Valid Process Support Tool for Remote Negotiation", January, 1998.
- 426. Pawan S. Budhwar and Harish C. Jain, "Evaluating Levels of Strategic Integration and Development of Human Resource Management in Britain", March, 1998.
- 427. Halit Üster and Robert F. Love, "Application of Weighted Sums of Order p to Distance Estimation", April, 1998.



Junis REF HB 74,5 , R47 NO.428

ł

J

4