UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dpto. de Matemática Aplicada

Extensión de funciones $\phi-$Lipschitz y aplicaciones

Trabajo Fin de Máster

Máster Universitario en Investigación Matemática

AUTOR/A: González Cortés, Álvaro

Tutor/a: Calabuig Rodriguez, Jose Manuel

Cotutor/a: Sánchez Pérez, Enrique Alfonso

Director/a Experimental: ARNAU NOTARI, ANDRES ROGER

CURSO ACADÉMICO: 2022/2023

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Master's thesis - Course 2022/ 2023

Extensions of ϕ-Lipschitz functions and applications

Author: **Álvaro González Cortés**

Diector: José Manuel Calabuig Rodríguez

Codirectors: Andrés Roger Arnau Notari
Enrique Alfonso Sánchez Pérez

Abstract

The classical Lipschitz real function extension theorems, due to McShane and Whitney, have found numerous applications in many fields, such as economics, mathematical analysis, and recently, in the field of artificial intelligence. These theorems can be generalized in various directions, extending the class of functions to which they can be applied, or weakening the metric conditions. In all cases, it allows to extend functions defined on metric subspaces to the whole space, preserving the Lipschitz constant.

In this work an increasing, positive, subadditive function ϕ is introduced which, when composed with the metric, gives another function with properties similar to the original metric. In the resulting space, Katetov-type functions can be defined from the same metric function, which are Lipschitz and which can also satisfy certain additional conditions. Recently, and with the intention of providing a functional basis for the definition of numerical indices in different disciplines (economics, foresight, demography, etc.), the notion of index space has been introduced. The indices are real Lipschitz functions, the referred Katetov functions being canonical examples.

The results of this work generalise the already known results on Lipschitz indices for the case of the ϕ-Lipschitz, besides the study of the compactness of the set of the corresponding standard indices. The properties of the approximation that make it possible to work with this functional basis for the design of artificial intelligence algorithms on ϕ-metric models will also be presented.

Resumen

Los teoremas clásicos de extensión de funciones reales de Lipschitz, debidos a McShane y Whitney, presentan numerosas aplicaciones en varias áreas, tales como la economía, el análisis matemático y, más recientemente, en el campo de la inteligencia artificial. Estos teoremas se pueden generalizar en varias direcciones, extendiendo la clase de funciones en las cuales pueden ser aplicados, o debilitando las condiciones métricas. En todos estos casos, esto permite extender funciones definidas en subespacios métricos al espacio entero, preservando la constante de Lipschitz.

En este trabajo se introduce una función ϕ creciente, positiva y subaditiva que, al componerla con una métrica, se obtiene otra función con similares propiedades a la métrica original. En el espacio métrico resultante, funciones del tipo Katetov pueden ser definidas a partir de esta misma función distancia, que serán Lipschitz y que pueden satisfacer además condiciones adicionales. Recientemente, y con la intención de proveer una base funcional para la definición de índices numéricos en diferentes disciplinas (economía, prospectiva, demografía, etc.), la noción de espacio de índices es introducida. Estos índices son funciones reales de Lipschitz, siendo las mencionadas funciones de Katetov ejemplos canónicos.

Los resultados de este trabajo generalizan los ya conocidos sobre índices de Lipschitz para el caso de los ϕ-Lipschitz, además del estudio de la compacidad del conjunto de los correspondientes índices estándard. También se presentarán las propiedades de la aproximación que hacen posible trabajar con esta base funcional para el diseño de algoritmos de inteligencia artificial en modelos ϕ-métricos.

Contents

Abstract **I**

Resumen **III**

Introduction **1**

1 Lipschitz maps and ϕ functions **3**

1.1 First definitions and examples 3

1.2 ϕ-Lipschitz maps 6

2 Index spaces **13**

2.1 Introduction to index spaces 13

2.2 Extensions of general indices 14

2.2.1 Approximation through standard indices 15

2.2.2 Extension theorems: McShane and Whitney formulas 18

3 Applications **21**

3.1 Implementation 21

3.2 Testing the algorithm 24

3.2.1 AARP livability index 24

3.2.2 Further analysis 26

Conclusions **31**

Algorithms **33**

Bibliography **44**

Introduction

Lipschitz's functions were defined for the first time to study the existence of solutions of differential equations, more specifically for Cauchy problems, the Picard-Lindelöf theorem [1, p.184-185] being the best known in this respect. Since then, these kinds of functions have been considered in other areas such as optimisation [2], machine learning [3] or manifolds [4]. This type of functions refers to the concept of distance, as it relates the distance of the images to the distance of the points in the domain, so they are defined in metric spaces.

In this work, we are interested in a classical result about how Lipschitz functions can be defined beyond their domain while preserving their condition. This question has many practical applications, as we can estimate unknown values of the function while preserving the original relation. An example of this can be found in index theory. We can understand an index as the evaluation of certain elements through the compilation of several indicators and evaluation of meaningfull factors, providing a single value that is significant and that allows comparisons or rankings. Canonical examples of indices are stock market indices, developed to represent the state of the financial markets or economy [5] such as the American S&P 500 or the Spanish IBEX35, or also in the social sciences, where university indices and rankings have gained notoriety [6]. However, the task of analysing all elements of interest may be difficult or impossible to implement in practice. This is where the extension of indices from known values comes into consideration.

The purpose of our work is to introduce a new type of Lipschitz-based functions to generalise this concept and to obtain better properties from the extension point of view. We will study the properties and examples of this function space related to the objective of our interest, we will present both its theoretical framework and the area on which we want to apply it (index theory) and we will analyse the performance and compare it to already existing procedures. For this purpose, we have structured this work as follows. In Chapter 1 we will present a detailed introduction to the Lipschitz condition and metric spaces, which are the basis of the models we will consider. After that, we will focus on motivating and developing the generalisation proposal we have introduced, analysing the behaviour of this new condition with respect to the main points of interest of the original one, such as extension results. Chapter 2 will focus on the mathematical framework on which the indices and the corresponding rankings are based. We will study index spaces based on the new class of functions introduced, and we will present two techniques that allow us to extrapolate the information provided by the indices for elements in which it is not defined. Thus, we will provide possible approximations for these elements, which will be the main point of interest. Then, in Chapter 3, we will see how to apply the results seen so far in algorithms that numerically implement the extension processes described. We will detail the examples in which we study their operation, thus providing the context in which they can be found, and special emphasis will be placed on the differences in performance with respect to the original alternative and to other procedures that can also be considered. Finally, in the Conclusions section, we will summarise the main aspects dealt with throughout this work and indicate possible lines of development that can be pursued in the future.

All the algorithms and visualizations of this Master's Thesis can be found in the link of GitHub https://github.com/Algoncor/TFM.

Chapter 1

Lipschitz maps and ϕ functions

We will begin this chapter by recalling results and basic concepts related to Lipschitz maps, previously establishing the theoretical framework in which they are found: metric spaces. The work of Deza et al. [7] and Cobzaş et al. [8] have been followed for this part. We will then focus on the main objective of this chapter: to generalise the concept of Lipschitz maps by constructing a larger family of maps while keeping their most interesting properties. For this purpose, a certain class of functions will also be introduced and studied, with special emphasis on their relation to metric spaces.

1.1 First definitions and examples

The following definition formalises the intuitive concept we may have of distance. We can interpret a distance d as a measure of how close two elements of a set are to each other, being 0 if they are the same element and larger values of the distance correspond to more distant elements.

Definition 1.1. Let D a nonempty set and $d\colon D \times D \to \mathbb{R}^+$, where $\mathbb{R}^+$ is de set of non-negative real numbers. It is said that d is a *distance* or *metric* if it satisfies the following conditions:

(i) $d(a,b) = 0$ if and only if $a = b$ (identity of indiscernibles),

(ii) $d(a,b) = d(b,a)$ for all $a, b \in D$ (symmetry),

(iii) $d(a,b) \leq d(a,c) + d(c,b)$ for all $a, b, c \in D$ (triangle innequality).

It follows from these three conditions that $d(a,b) \geq 0$ for all $a, b \in D$, and therefore d is well-defined. The pair (D,d) is called *metric space*, and for any subset $D_0 \subseteq D$ with the distance restricted to it, $d\restriction_{D_0}$, is called *metric subspace*. Moreover, it is said that D_0 is *bounded* if there exists $M > 0$ such that $d(x,y) \leq M$ for all $x, y \in D_0$. From now on we will assume that (D,d) denotes a metric space and $(D_0, d\restriction_{D_0})$ a metric subspace of it.

Example 1.2. Some examples of metrics will be presented now for illustrative purposes.

- Given a set D, the map defined by $d(x, y) = 1$ if $x \neq y$ and $d(x, x) = 0$ is a metric, know as *discrete metric*.

- Every normed space $(X, \|\cdot\|)$ is a metric space considering $d(x, y) = \|x - y\|$, which verifies the requirements to be a metric. Some examples are the canonical distance in $\mathbb{R}$, which is $d(x, y) = |x - y|$ for $x, y \in \mathbb{R}$, and the *integral metric* in $C_{[a,b]}$, the set of real (or complex) continuous functions defined in the segment $[a, b]$. The canonical norm in this vector space is
$$\|f\| = \int_a^b |f(x)|\, dx,$$
so we can define a metric as $d(f, g) = \displaystyle\int_a^b |f(x) - g(x)|\, dx$.

- Let A a nonempty set, $n \in \mathbb{N}$ and $D = A^n$. The *Hamming distance* in D is defined for $a = (a_1, \dots, a_n),\ b = (b_1, \dots, b_n) \in D$ as the number of $k \in \{1, \dots, n\}$ such that $a_k \neq b_k$. For instance, if A is the set of letters of the Latin alphabet and $n = 5$, taking $a = fight$ and $b = night$, $d(a, b) = 1$ because only the first letter does not match, but if $c = write$ then $d(a, c) = 5$ because no one match.

- Given a set V and $E \subseteq V \times V$ a set of unordered pairs of elements of V, the pair $G = (E, V)$ is a *graph*. If $u, v \in V$, a $(u - v)$ path is a sequence $(w_0, w_1), (w_1, w_2), \dots, (w_{n-1}, w_n) \in E$ such that $w_0 = u$, $w_n = v$ and $w_i \neq w_j$ if $i \neq j$. The length of a path is the number of elements in the sequence, and a graph is called *connected* if for every $u, v \in V$ there exists a $(u - v)$ path with finite length. In this context, in a connected graph G we can define a metric in V for all $u, v \in V$ as the shortest leght of all $(u - v)$ paths.

More examples can be found in [7].

Each metric space (D, d) is also a topological space if we consider the topology induced by the distance, in which the basis neighbourhoods of $x \in D$ are the open balls $B_\varepsilon(x) = \{y \in D : d(x, y) < \varepsilon\}$, being $\varepsilon > 0$.

A property studied in topological spaces is compactness, that seeks to generalize the notion of a closed and bounded subset of Euclidean space. The definition of a compact topological space can be found in [7], but here we present a characterization of this property in the case of metric spaces, since we will be interested in this property later on.

Proposition 1.3. (D, d) is a compact metric space if, and only if, for each sequence $\{a_n\}_n \subseteq D$ there exists $\{a_{n_k}\}_k$ a subsequence of $\{a_n\}_n$ that converge to some $a \in D$. That is, there exists $a \in D$ such that $\lim_k d(a_{n_k}, a) = 0$.

Let us introduce the main concept of this section, which will be fundamental to the rest of the work: the Lipschitz maps. In these maps we find a relationship between the distance of the images of two points and the distance between the points themselves. Intuitively, a map is Lipschitz if the "slope" between any two points can be bounded by the same constant. Formally, the definition is as follows.

Definition 1.4. Let (D, d) and (R, r) metric spaces. A map $f\colon D \to R$ is called *Lipschitz* if there exists a constant $L \in \mathbb{R}$ such that

$$r(f(x), f(y)) \leq Ld(x, y), \text{ for all } x, y \in D.$$

It is said that L is the *Lipschitz constant* of f and we say f is L-Lipschitz to emphasize this constant. The infimum of all constants satisfying last inequality is the *Lipschitz norm* of f, denoted by $Lip(f)$. This can be written as

$$Lip(f) = \sup_{x,y\in D} \frac{r(f(x), f(y))}{d(x, y)}.$$

Example 1.5. Let us present some examples of Lipschitz maps:

- Consider the metric subspace $\mathbb{R}^+$ of $\mathbb{R}$ with their canonical distance. The function $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ defined by $f(x) = \log(x + 1)$ is 1-Lipschitz. Indeed, let $x, y \in \mathbb{R}^+$ and suppose, without loss of generality, that $y \geq x$. Since $\log(x + 1) \leq x$ for $x \geq 0$ we get

$$\begin{aligned} f(y) - f(x) &= \log(y + 1) - \log(x + 1) \\ &= \log\left(\frac{y+1}{x+1}\right) = \log\left(\frac{y-x}{x+1} + 1\right) \\ &\leq \frac{y-x}{x+1} \leq y - x, \end{aligned}$$

  where last inequality holds because $x + 1 \geq 1$. From this we have that f is Lipschitz with $Lip(f) \geq 1$. Moreover, we have also that

$$\frac{f(y) - f(x)}{y - x} \leq 1,$$

  so $Lip(f) \leq 1$ and f is 1-Lipschitz.

- Let $D = \mathbb{R}^2$ and $d(x, y) = \|x - y\|_2$, being $\|(x_1, x_2)\|_2^2 = x_1^2 + x_2^2$. If $S^1 = \{x \in D : \|x\|_2 = 1\}$, the map $f\colon S^1 \to S^1$ defined by $f(x_1, x_2) = (x_2, x_1)$ is 1-Lipschitz.

- Consider $l^\infty(\mathbb{C})$ the space of bounded sequences in $\mathbb{C}$, that is, $\mathbb{C} \supset \{x_n\}_n \in l^\infty(\mathbb{C})$ if $\sup_n |x_n| < \infty$. In this space we consider the metric induced by their canonical norm $\|x\|_\infty = \sup\limits_n |x_n|$. The map $B\colon l^\infty(\mathbb{C}) \to l^\infty(\mathbb{C})$ defined by $B(x_1, x_2, \ldots, x_n, \ldots) = (x_2, x_3, \ldots, x_n \ldots)$, know as *backward shift operator*, is 1-Lipschitz.

Let us suppose now that $f\colon S \subseteq X \to \mathbb{R}$ is a function defined on a subset S of X. Several classical problems consist of defining f in $X \setminus S$ while maintaining some initial property of the map f. This type of problems, known as extension problems, are present in several situations, although in some of them the explicit expression cannot be obtained. The property we start from in our work is the Lipschitz condition, due to the useful extensions formulas that can be obtained, and the following theorem guarantees it.

Theorem 1.6. If $f\colon D_0 \subseteq D \to \mathbb{R}$ is an L-Lipschitz function, then there exists an L-Lipschitz function $F\colon D \to \mathbb{R}$ such that $F\restriction_{D_0}= f$.

This function F is not unique, and two possible formulas are

$$F^M(x) := \sup_{y\in D_0}\{f(y) - Ld(x,y)\} \quad \text{and} \quad F^W(x) := \inf_{y\in D_0}\{f(y) + Ld(x,y)\},$$

which are known as the McShane and Whitney extensions respectively.

Remark 1.7. We notice that any extension F of a Lipschitz function f verifies that $F^M \leq F \leq F^W$. Moreover, for any $t \in (0,1)$, $F := tF^W + (1-t)F^M$ is also a Lipschitz extension of f with the same constant.

The proof of these results can be found in [8].

1.2 ϕ-Lipschitz maps

Our attention will now turn to whether we can generalise the Lipschitz condition, so that we can still have extension theorems for a new class of functions satisfying a new, more relaxed condition. In this work we will give an affirmative answer by introducing the ϕ-Lipschitz functions. For this purpose, we will first introduce the class of functions Φ.

Definition 1.8. We will say that $\phi \in \Phi$, being $\phi\colon \mathbb{R}^+ \to \mathbb{R}^+$, if for each $x, y \in \mathbb{R}^+$ it holds that

(i) $\phi(x+y) \leq \phi(x) + \phi(y)$ (subadditivity),

(ii) $\phi(x) < \phi(y)$ if $x < y$ (strictly monotonically increasing),

(iii) $\phi(0) = 0$,

(iv) ϕ is continuous in $\mathbb{R}^+$.

Let us remark that conditions (ii) and (iii) guarantee that $\phi(x) \geq 0$ for all $x \in \mathbb{R}^+$, so these functions are well defined.

Remark 1.9. Condition (iv) can in fact be reduced to the requirement that ϕ is right-continuous at 0, since from this property and the other ones can be deduced the global continuity. Let us see this. Take $x > 0$ and note that, by (i) and (ii), holds

$$\lim_{h \to 0^-} \phi(x+h) \leq \lim_{h \to 0^+} \phi(x+h) \leq \lim_{h \to 0^+} \phi(x) + \phi(h) \leq \phi(x), \tag{1.1}$$

being last inequality consequence of (iii) and the right-continuity at 0. On the other hand, $\phi(x) \leq \phi(x-h) + \phi(h)$ for $h > 0$ by (iii), so

$$\phi(x) \leq \lim_{h \to 0^+} \phi(x-h) + \phi(h) = \lim_{h \to 0^-} \phi(x+h). \tag{1.2}$$

Putting (1.1) and (1.2) together we conclude that $\phi(x) = \lim_{h \to 0^-} \phi(x+h)$, that is, ϕ is left-continous at x. By similar arguments it can be shown that ϕ is also right-continuous at x, so ϕ is globally continuous.

Having made this presentation, we can now give the definition of ϕ-Lipschitz maps.

Definition 1.10. Let (D, d) and (R, r) be metric spaces and $f\colon D \to R$ a map. For $\phi \in \Phi$ we will say that f is a *ϕ-Lipschitz map* if there exists $K > 0$ such that

$$r(f(x), f(y)) \leq K\phi(d(x, y)), \text{ for all } x, y \in D.$$

From now on, when we will compose the function $\phi \in \Phi$ with the metric d we will write d_ϕ. That is, $d_\phi = \phi \circ d$.

We notice that any L-Lipschitz map is a ϕ-Lipschitz map for $\phi(x) = x$ and $K = L$. Moreover, the function $f\colon \mathbb{R}^+ \to \mathbb{R}^+$ defined by $f(x) = \sqrt{x}$ is not Lipschitz but it is ϕ-Lipschitz for $\phi = f \in \Phi$ and $K = 1$. So effectively this new class of maps generalise the original L-Lipschitz class of maps.

In addition, the boundary of the distance between the images of two points provided by the Lipschitz condition can be improved if it is considered as ϕ-Lipschitz. This case, which as we will see in Chapter 3 is fundamental in practical matters, is presented for example by the function $f(x) = \log(x+1)$. We know that f is 1-Lipschitz and so $d(f(x), f(y)) \leq d(x, y)$, but f is also ϕ-Lipschitz for $\phi = f$ and $K = 1$, and therefore $d(f(x), f(y)) \leq \log(d(x, y) + 1)$, which is a better bound because $\log(d(x, y) + 1) \leq d(x, y)$.

Actually, our proposal consists in redefining the distance in question so that new maps that meet the proposed condition (in addition to the Lipschitz maps) appear in the resulting metric space. In this way, by varying the ϕ one has direct control over how to change the original metric so that the resulting better suits the problem at hand. Next we will see that d_ϕ does indeed define a distance.

Proposition 1.11. Let (D, d) a metric space and $\phi \in \Phi$. Then d_ϕ is a metric.

Proof. Since ϕ is an injective function (because it is strictly monotonic) and $\phi(0) = 0$, it is clear that $d_\phi(x, y) = 0$ if, and only if, $d(x, y) = 0$. And since d is a metric this happens if, and only if, $x = y$, and we have just proved the identity of the indiscernibles for d_ϕ.

From the symmetry of d as metric we deduce the symmetry of d_ϕ.

Finally, for all $x, y, z \in D$, the triangle inequality for d implies $d(x, z) \leq d(x, y) + d(y, z)$. Take into account the monotonicity and subadditivity of ϕ we get

$$d_\phi(x, z) \leq \phi(d(x, y) + d(y, x)) \leq d_\phi(x, y) + d_\phi(y, z),$$

so d_ϕ verifies the triangle inequality too. $\square$

The reciprocal of this result is false in general. For example, if $D = \mathbb{R}$ and we let $\phi \in \Phi$ as $\phi(x) = \sqrt{x}$ for $x \geq 0$, and $d(x, y) = |x - y|^2$, we know that $d_\phi = |x - y|$ is a metric on D but d is not. To see this it is enough to take $x = 0$, $y = 1$ and $z = 3$ and check that $d(x, z) > d(x, y) + d(y, z)$, and therefore d does not satisfy the triangular inequality.

On the other hand, from the assumption that for each (D, d) metric space d_ϕ is a metric, we can deduce some conditions that ϕ must satisfy. For example, it is easy to check that $\phi(0) = 0$, and we will see next that ϕ must be subadditive. Let $x, y \in \mathbb{R}^+$ and define a metric space $D = \{a, b, c\}$ such that $d(a, b) = x, d(b, c) = y$ and $d(a, c) = x + y$. As d_ϕ is a metric, by the triangle inequality we obtain $\phi(x + y) = d_\phi(a, c) \leq d_\phi(a, b) + d_\phi(b, c) = \phi(x) + \phi(y)$, and so ϕ is subadditive. However, the condition of monotonicity and continuity that ϕ must satisfy may not be met, that is, d_ϕ can be a metric for all metric space (D, d) without ϕ being strictly monotone increasing neither continuous. For example, consider ϕ defined as $\phi(x) = 0$ if $x = 0$ and $\phi(x) = 1$ if $x > 0$. For each (D, d) metric space we note that:

(I) For $a, b \in D$, condition $d_\phi(a, b) = 0$ holds if, and only if, $d(a, b) = 0$ because ϕ is only null at 0. Since d is a metric we have that it holds if, and only if, $a = b$, so d_ϕ satisfy the identity of indiscernibles.

(II) Since $d(a, b) = d(b, a)$ for all $a, b \in D$ it is clear that $d_\phi(a, b) = d_\phi(b, a)$, that is, d_ϕ is symmetric.

(III) Take $a, b, c \in D$ and write $x = d(a, b), y = d(b, c)$ and $z = d(a, c)$. If a, b, c are different from each other then $x, y, z > 0$ since d is a metric, and so $\phi(x) = \phi(y) = \phi(z) = 1$. From this we conclude that d_ϕ satisfy the triangular innequality for a, b, c. In case some of the elements of D that we have taken coincide, the triangular innequality of d_ϕ for them is trivial.

We have just proved that d_ϕ is a metric for any metric space, and yet ϕ is not strictly monotonic nor continuous.

For a better understanding of how d and d_ϕ relate to each other, we present in Figures 1.1 and 1.2 a graphical representation of the behaviour of two concrete metrics: the usual distance d in $\mathbb{R}$ and its composition with $\phi(x) = \log(1+x)$, that is, d_ϕ. In Figure 1.1 one can observe that d_ϕ "smoothes" the distance between two real numbers respect d. More specifically, d and d_ϕ have a similar behaviour for x and y that are "close" to each other, but for more distant values d_ϕ attenuates growth with respect to d. In Figure 1.2 a comparison is provided between the behaviour of triangular inequality of d and d_ϕ. As before, we can see how the logarithmic growth of ϕ carries over to d_ϕ, and also over to their triangular innequality.

Example 1.12. More examples of ϕ-Lipschitz maps are presented below.

- A map $f\colon \mathbb{R}^n \to \mathbb{R}$ is called *α-Hölder continuous* if there exists $C > 0$ and $\alpha > 0$ such that
$$r(f(x), f(y)) \leq C d(x,y)^\alpha, \quad \text{for all } x, y \in D.$$
These maps are ϕ-Lipschitz maps for $K = C$. Indeed, if $\alpha \in (0,1]$ then ϕ defined as $\phi(x) = x^\alpha$ belongs to Φ and f is ϕ-Lipschitz for $K = C$. If $\alpha > 1$ it can be proven that f is a constant map, so it is ϕ-Lipschitz for any $\alpha \geq 0$.

- Consider $\mathbb{R}$ equipped with its usual metric and let $\phi \in \Phi$. From the subadditivity it follows that $\phi(x) - \phi(y) \leq \phi(x-y)$ for $x \geq y$, so $d(\phi(x), \phi(y)) \leq \phi(d(x,y))$. That is, every $\phi \in \Phi$ is a ϕ-Lipschitz function for $K = 1$. Examples of ϕ functions, apart from those already presented, are $\phi(x) = \arctan(x)$, $\phi(x) = x(x+1)^{-1}$ or $\phi(x) = x(x^2+1)^{-1/2}$.

In the next result, we will see that the ϕ-Lipschitz functions can be extended similarly to the ones seen in Theorem 1.6. We will see that this is due to the conditions that the Φ functions must fulfil.

Theorem 1.13. Let (D, d) be a metric space. If $f\colon D_0 \subseteq D \to \mathbb{R}$ is a ϕ-Lipschitz function for $K > 0$, then there exists a ϕ-Lipschitz function $F : D \to \mathbb{R}$ for K such that $F \restriction_{D_0} = f$.

This function F is not unique, and two possible formulas are

$$F^M(x) := \sup_{y \in D_0} \{f(y) - K d_\phi(x,y)\} \quad \text{and} \quad F^W(x) := \inf_{y \in D_0} \{f(y) + K d_\phi(x,y)\},$$

which we will call McShane and Whitney extensions respectively, as in the original theorem.

Proof. The techniques used in this proof are similar to those used in Theorem 1.6, but for the sake of completeness we will bring them here. Furthermore, it is enough to show that one of the two functions above satisfies the result. We will do the proof for F^W, for F^M is analogous.

Given $x \in D_0$ let us see that $F^W(x) = f(x)$. On the one hand, $f(x) \leq f(y) + K d_\phi(x,y)$ for all $y \in D_0$ because f is ϕ-Lipschitz in D_0, so

$$f(x) \leq \inf_{y \in D_0} \{f(y) + K d_\phi(x,y)\} = F^W(x). \tag{1.3}$$

On the other hand, since $x \in D_0$,

$$F^W(x) = \inf_{y \in D_0} \{f(y) + K d_\phi(x, y)\} \leq f(x) + K d_\phi(x, x) = f(x), \tag{1.4}$$

so, by puting together (1.3) and (1.4), we conclude $F^W(x) = f(x)$ for an arbitrary $x \in D_0$, so $F^W \restriction_{D_0} = f$.

In order to show that F^W is ϕ-Lipschitz for K, let $x, z \in D$ and suppose, without loss of generality, that $F^W(x) - F^W(z) \geq 0$. Given $\varepsilon > 0$ there exists $y_\varepsilon \in D_0$ such that

$$\varepsilon - f(y_\varepsilon) - K d_\phi(z, y_\varepsilon) > \sup_{y \in D_0} \{-f(y) - d_\phi(z, y)\}. \tag{1.5}$$

Moreover,

$$\inf_{y \in D_0} \{f(y) + K d_\phi(x, y)\} \leq f(y_\varepsilon) + K d_\phi(x, y_\varepsilon). \tag{1.6}$$

Taking into account (1.5) and (1.6) we have

$$\begin{aligned} F^W(x) - F^W(z) &= \inf_{y \in D_0} \{f(y) + K d_\phi(x, y)\} + \sup_{y \in D_0} \{-f(y) - K d_\phi(z, y)\} \\ &< \varepsilon - f(y_\varepsilon) - K d_\phi(z, y_\varepsilon) + f(y_\varepsilon) + K d_\phi(x, y_\varepsilon) \\ &\leq \varepsilon + K d_\phi(x, z) + K d_\phi(z, y_\varepsilon) - K d_\phi(z, y_\varepsilon) = \varepsilon + K d_\phi(x, z). \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we conclude that $F^W(x) - F^W(z) \leq K d_\phi(x, z)$. □

One of the aims of the next section will be to establish compactness results for subsets of certain ϕ-Lipschitz functions. For this, we need to establish how the compactness of a metric space is related to ϕ-functions. The following result will focus on this question.

Proposition 1.14. If (D, d) is compact, then (D, d_ϕ) is compact too.

Proof. Let $\{a_n\}_n$ be a sequence in D. As D is compact we know that there exists $\{a_{n_k}\}_k$ subsequence of $\{a_n\}_n$ that converges to $a \in D$. That is, $\lim_k d(a_{n_k}, a) = 0$. Since ϕ is a continuous function we get

$$0 = \phi(0) = \phi\left(\lim_k d(a_{n_k}, a)\right) = \lim_k d_\phi(a_{n_k}, a),$$

so $\{a_{n_k}\}_k$ converges to $a \in D$, and by Proposition 1.3 we conclude that (D, d_ϕ) is a compact metric space. □

Remark 1.15. We note that to prove 1.13 we have made use of conditions (i)-(iii) of Φ, being condition (iv) only necessary to prove the latter result. These conditions that $\phi \in \Phi$ must satisfy are known in the literature as *modules of continuity*. These are increasing functions $\omega\colon [0, +\infty) \to [0, +\infty)$ that cancels at 0 and are continuous at 0, but depending on the context they may satisfy more conditions.

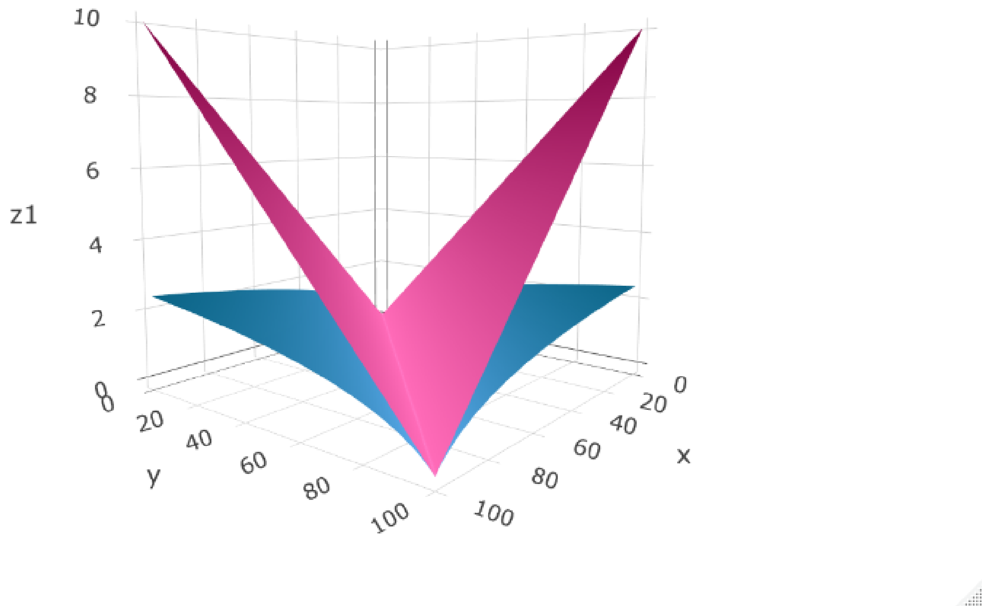
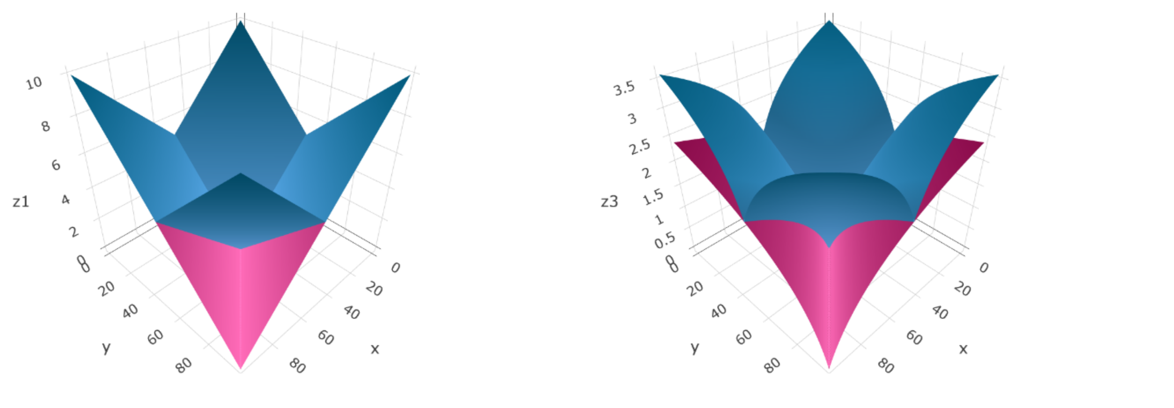


Figure 1.1: Comparative between $d(x, y)$, in pink, and $d_\phi(x, y)$, in blue.



(a) Comparison between $d(x, y)$, in pink, and $d(x, 0) + d(0, y)$, in blue.

(b) Comparison between $d_\phi(x, y)$, in pink, and $d_\phi(x, 0) + d_\phi(0, y)$, in blue.

Figure 1.2: Representation of the triangular inequality of d and d_ϕ

Chapter 2

Index spaces

This chapter will focus first on introducing the concept of index space and presenting definitions and related concepts necessary for the rest of the chapter. Later on, we will look at the index extension, which will be the main focus of this chapter. For this purpose, we will present two different techniques: one of them is based on two classical function extension formulas and the other on approximations of what we will call standard indices. For this chapter we have followed the work of Erdoğan et al. [9].

2.1 Introduction to index spaces

The models we are interested in studying start from a metric space (D, d) in which an index I is defined. This index is a ϕ-Lipschitz function $I \colon D \to \mathbb{R}$ that provides a meaningful value on the elements of D, depending on the model in question. For this we assume that the index I is controlled by the distance d, that is, I represents a quantity whose properties are implicitly represented by d. That is why we assume that I is a ϕ-Lipschitz function, and with it, for example, a ranking of the elements of D can be constructed according to the value of its index. In addition, we will impose other conditions on the index related to d, in order to achieve a better structure to work with. In this context we will say that the triplet (D, d, I) is an index space, and from now on we will suppose that the metric is bounded, that is, $\operatorname{diam}(D) = \sup_{a,b\in D} d(a, b) < \infty$. In the following we will look at all these definitions that will be needed in the work.

Definition 2.1. Let (D, d) be a metric space. An index $I \colon D \to \mathbb{R}$ is *C-bounded*, for $C > 0$, if $\sup_{a\in D} |I(a)| \leq C$. The infimum of all constants C satisfying the last inequality is known as *boundedness constant* of I, and is denoted by $B(I)$. That is, $B(I) = \sup\limits_{a\in D} |I(a)|$.

The following normalisation property for a constant Q will be useful for further comparison between several indexes. This condition, which relates I to d, will be complementary to the ϕ-Lipschitz condition.

Definition 2.2. Let (D, d) be a metric space. An index $I: D \to \mathbb{R}$ is *Q-normalized*, for $Q > 0$, if

$$d_\phi(a, b) \leq Q(|I(a)| + |I(b)|), \quad \text{for all } a, b \in D.$$

The functions that satisfy this condition are kwow in literature as *Katetov functions*. The infimum of all constants Q satisfying last inequality is know as *normalization constant* of I, and is denoted by $N(I)$. Moreover, we note that if I is bounded and Q-normalized then $d(D, D) \leq 2QB(I)$.

Remark 2.3. The condition that I is Q-normalised, that is, $N(I) < \infty$, implies that I can only nullify at one point. Indeed, given $a, b \in D$ such that $I(a) = I(b) = 0$, then $d(a, b) \leq Q(I(a) + I(b)) = 0$, and so $a = b$. As we will see later, this will mean that there can only be one "optimal" element in the space, in the sense that I can only be 0 at a single point.

Definition 2.4. Let (D, d) be a metric space. An index $I: D \to \mathbb{R}$ is *ϕ-coherent*, for $\phi \in \Phi$ and $K > 0$, if

$$|I(a) - I(b)| \leq K d_\phi(a, b), \quad \text{for all } a, b \in D.$$

That is, I is ϕ-Lipschitz for K. The infimum of all constants K satisfying last inequality is the *coherence constant* of I, and we will write it $C(I)$.

Note that, if $R := \inf_{a \in D} I(a) \in \mathbb{R}^+$, we have

$$B(I) - R = \sup_{a \in D} I(a) - R \leq \sup_{a,b \in D} |I(a) - I(b)| \leq C(I)\phi(d(D, D)).$$

Thus, $B(I) \leq C(I)\phi(d(D, D)) + R$, which means that every ϕ-Lipschitz map in an index space is bounded. Furthermore, if $N(I) < \infty$ we get

$$\phi(d(D, D)) \leq \phi(2N(I)B(I)) \leq 2\phi(N(I)B(I)),$$

so we can give an estimate of the boundedness constant only from properties of the index (coherence and normalization constants and their infimum).

2.2 Extensions of general indices

When dealing with practical issues in index spaces, the index value will not always be available for all elements of the metric space. This may be the case, for example, when it is costly or impossible to obtain all the data in the model. It is in this context that it is useful to have tools that allow us to approximate the index for unindexed elements. In this section, we propose two ways of doing this: the first way is to identify/approximate the index of interest using what we will call standard indices, while the second way is to approach the issue as a Lipschitz regression problem.

2.2.1 Approximation through standard indices

The structure of a metric space (D, d) provides a behaviour between the elements of D that will somehow also be transferred to the index I. For example, if two elements $a, b \in D$ are "close" to each other (that is, $d(a, b)$ is a value close to 0), the value of the index at those points must be similar. This is why the metric space provides standard indices related to individual points $a \in D$, which we denote by I_a. This index can be defined, after choosing a reference point $a \in D$, as $I_a(b) = d(a, b)$ for each $b \in D$. Typically, the criterion for choosing a is based on finding the element that minimises a given property, although any other element can also be chosen. Moreover, we can also consider a function $\phi \in \Phi$ to improve the properties of the metric by taking $I_a(b) = d_\phi(a, b)$. In the following proposition we will see why an index defined in this way can be considered *standard*.

Proposition 2.5. An index defined by $I_a(\cdot) = d_\phi(a, \cdot)$ for some $a \in D$ is 1-normalized and ϕ-coherent for $K = 1$.

Proof. Given $b, c \in D$ note that $d_\phi(b, c) \leq d_\phi(b, a) + d_\phi(a, c) = I_a(b) + I_a(c)$, and so $N(I) \leq 1$. Moreover, $d_\phi(a, b) = I_a(b) + I_a(a)$ and then $N(I) = 1$, so I_a is 1-normalized. On the other hand, by the triangle inequality of d_ϕ it is clear that $|I_a(b) - I_a(c)| = |d_\phi(a, b) - d_\phi(a, c)| \leq d_\phi(b, c)$. In addition, note that $|I_a(a) - I_a(b)| = I_a(b) = d_\phi(a, b)$, and therefore I_a is ϕ-Lipschitz for $K = 1$. □

From this result it follows that from any metric space and for any point a in it, we can construct a standard index centred on it. These indices satisfy the properties of normalisation and coherence that we have seen, and they also nullify at a and $I(b) > 0$ for all $a \neq b \in D$. In the following result we will see that these properties in fact characterise the standard indices.

Proposition 2.6. Let $I : D \to \mathbb{R}$ be a 1-normalized and ϕ-coherent index for $K = 1$. If there exists $a \in D$ such that $I(a) = 0$, then $I = I_a$.

Proof. For each $b \in D$, as I is ϕ-coherent for $K = 1$ we have that $I(b) = I(b) - I(a) \leq |I(b) - I(a)| \leq d_\phi(a, b)$. Moreover, note that $d_\phi(a, b) \leq I(a) + I(b) = I(b)$ since I is 1-normalized. It follows from the above two inequalities that $I(b) = d_\phi(a, b)$ and so $I = I_a$. □

Our aim from now on will be to study how we can approximate any given index by means of standard indices. The idea is to show that general indices can be approximated by limits and translations of standard indices. Thus, the natural extension of the class of standard indices is given by considering its closure. So, in the rest of the section we will study the structure and results of compactness in index spaces in order to have the necessary theoretical tools, and then we will focus on the main issue of the approximation.

For $C > 0$ consider the space $\mathcal{F}_C := \{I : D \to \mathbb{R} : B(I) \leq C\}$. In it we can define two natural topologies: the *uniform topology* with norm $B(\cdot)$ and the *pointwise convergence topology*. In the first one the basic neighbourhoods are $V_\varepsilon(I_0) = \{I \in \mathcal{F}_C : B(I - I_0) < \varepsilon\}$, for $\varepsilon > 0$ and

$I_0 \in \mathcal{F}_C$. In the second one they are $V_{\varepsilon,a_1,\ldots,a_n}(I_0) = \{I \in \mathcal{F}_C : |I(a_i) - I_0(a_i)| < \varepsilon, i = 1, \ldots, n\}$ for $\varepsilon > 0$, $I_0 \in \mathcal{F}_C$ and $a_1 \ldots, a_n \in D$. To implement the approximation tool we will propose, we need to establish compactness results for certain subspaces of $\mathcal{F}_C$. To do so, we will refer to Tychonoff's theorem, a basic topological result on the compactness of the quotient topology.

Proposition 2.7. The following subespaces of $\mathcal{F}_C$ are compact with respect to the topology of pointwise convergence:

(i) $\mathcal{F}_C^0 := \{I \in \mathcal{F}_C : I \geq 0\}$,

(ii) $\mathcal{F}_C^1 := \{I \in \mathcal{F}_C^0 : |I(a) - I(b)| \leq K d_\phi(a, b),\ a, b \in D\}$, for $K > 0$,

(iii) $\mathcal{F}_C^2 := \{I \in \mathcal{F}_C^1 : R + d(a, b) \leq Q(I(a) + I(b)),\ a, b \in D\}$, for $Q > 0$ and $R \geq 0$.

Proof. If $\{I_\eta\}_{\eta\in\Lambda}$ is a net on $\mathcal{F}_C$ that converges pointwise, then $\lim_\eta I_\eta(a) \in [-C, C]$ for all $a \in D$, so the limit is a function in $\mathcal{F}_C$ too. In particular it also holds for $\mathcal{F}_C^0$ since $I_\eta(a) \in [0, C]$. In addition, we can identify each function of $\{I_\eta\}_{\eta\in\Lambda}$ with an element of $\Pi_{a\in D}[-C, C]$ (or $\Pi_{a\in D}[0, C]$ in case of $\mathcal{F}_C^0$), which are product of compact spaces in the product topology. By Tychonoff's theorem we conclude the result of (i).

To prove (ii) let $a, b \in D$ and $\{I_\eta\}_{\eta\in\Lambda}$ a net that converges pointwise. Note that

$$|\lim_\eta I_\eta(a) - \lim_\eta I_\eta(b)| \leq K d_\phi(a, b)$$

since $I_\eta \in \mathcal{F}_C^1$ for each $\eta \in \Lambda$. So $I_\eta \in \mathcal{F}_C^1$ is closed, because limits of nets in this set is again a function in this set, and together with (i) we conclude (ii).

Note that $d(a, b) \leq Q(\lim_\eta I_\eta(a) + \lim_\eta I_\eta(b))$ for $a, b \in D$. The result of (iii) follows in the same way as the previous ones. $\square$

At this point we will explain in more detail our proposal for a standardised index approach. Given an index $I \in \mathcal{F}_C^2$, which is the class of indices discussed in this work, we are interested in finding a sequence of points in D whose associated standard indices "converge" to I. More precisely, the main result we will present in this section will establish a boundary between the index I and its infimum and a sequence of standard indices, associated to certain points of D. That is why we will actually work with the topology of pointwise convergence. However, in case of standard indices, we actually can get a more general result.

Proposition 2.8. Let (D, d) a compact metric space and consider the space of standard indices $S := \{d_\phi(a, \cdot) : a \in D\} \subset \mathcal{F}_C^2$. Then S is compact respect to the topology of uniform convergence.

Proof. Let $\{d_\phi(a_n, \cdot)\}_n \subset S$. As (D, d) is compact, by Proposition 1.14 we know that (D, d_ϕ) is also compact, so the sequence $\{a_n\}$ admits a subsequence $\{a_{n_k}\}_k$ convergent to $a_0 \in D$ respect

d_ϕ. So, for every $\varepsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that $d_\phi(a_{n_k}, a_0) < \varepsilon$ for all $k \geq k_0$. Then, by the triangular inequality we get

$$\sup_{b\in D} |d_\phi(a_{n_k}, b) - d_\phi(a_0.b)| \leq \sup_{b\in D} d_\phi(a_{n_k}, a_0) = d_\phi(a_{n_k}, a_0) < \varepsilon$$

for $k \geq k_0$. That is, $\{d_\phi(a_{n_k}, \cdot)\}$ converges uniformly to $d_\phi(a_0, \cdot)$, so by Proposition 1.3 we conclude the desired result. $\square$

We also need to name a type of sequence that will appear in the result we are looking for. The definition is as follows.

Definition 2.9. A sequence $\{a_n\}_n \subset D$ is *pointwise Cauchy* if for each $b \in D$ there exists $\lim_n d(a_n, b)$.

From our study of d_ϕ it is clear that this definition is equivalent to the next one: a sequence $\{a_n\}_n \subset D$ is *pointwise Cauchy* if for each $b \in D$ there exists $\lim_n d_\phi(a_n, b)$ for a given $\phi \in \Phi$. On the other hand, it is obvious that every convergent sequence is pointwise Cauchy, but the reciprocal is false. The sequence $a_n = 1/n$ in $D = (0, 1]$ equipped with their usual metric is a classic example of nonconvergent pointwise Cauchy sequence.

Theorem 2.10. For every $I \in \mathcal{R}_{C,K,Q}$, with

$$\mathcal{R}_{C,K,Q} := \{I \geq 0 : |I(a) - I(b)| \leq K d_\phi(a, b),\ \tfrac{1+KQ}{K} \inf(I) + d_\phi(a, b) \leq Q(I(a) + I(b)),\ B(I) \leq C\},$$

there exists a pointwise Cauchy sequence $\{a_n\}_n$ such that $I(b) \leq \inf(I) + \lim_n K d_\phi(a_n, b) \leq KQI(b)$ for each $b \in D$.

Proof. Let $b \in D$ and fix $n \in \mathbb{N}$. We know that there exists $a_n \in D$ such that $I(a_n) - \frac{1}{n} \leq \inf(I)$. Then

$$\begin{aligned}\inf(I) + K d_\phi(a_n, b) &\leq KQI(a_n) + KQI(b) - KQ \inf(I) \\ &\leq KQI(b) + KQI(a_n) - KQ\left(I(a_n) - \tfrac{1}{n}\right) = KQI(b) + \tfrac{KQ}{n}.\end{aligned}$$

In addition,

$$I(b) - I(a_n) \leq |I(b) - I(a_n)| \leq K d_\phi(a_n, b),$$

and therefore

$$I(b) \leq K d_\phi(a_n, b) + I(a_n) \leq K d_\phi(a_n, b) + \inf(I) + \tfrac{1}{n} \leq KQI(b) + \tfrac{1+KQ}{n} \tag{2.1}$$

for all $n \in \mathbb{N}$ and $b \in D$. Now note that $d_\phi(a_n, \cdot) = I_{a_n}(\cdot) \in S$, thus, from the compactness of S seen in Proposition 2.8 follows that there exists a subsequence $\{a_{n_k}\}_k$ such that $\lim_k d_\phi(a_{n_k}, b) = d_\phi(a_0, b)$ for each $b \in D$ and certain $a_0 \in D$. Therefore $\{a_{n_k}\}_k$ is a pointwise Cauchy sequence and from (2.1) we conclude that

$$I(b) \leq \inf(I) + K \lim_k d_\phi(a_{n_k}, b) \leq KQI(b)$$

for every $b \in D$. $\square$

This result leads us to identify $\tilde{I} := \inf(I) + K \lim_k d_\phi(a_{n_k}, \cdot)$ as an approximation of $I \in \mathcal{R}_{C,K,Q}$. In this case the maximum error committed would be

$$\sup_{b\in D} \left|\tilde{I}(b) - I(b)\right| \leq \sup_{b\in D} |KQI(b) - I(b)| = |KQ - 1|\, C.$$

So if $KQ \approx 1$ the comparison between $\tilde{I}$ and I is reasonable, but as these constants increase the approximation may deteriorate. In the particular case that $KQ = 1$ we would in fact have $\tilde{I} = I$. This situation holds for example if $K = Q = 1$, so that I would be a standard index as we saw in Proposition 2.6. If $a \in D$ is such that $I = I_a$, then $\inf(I) = I_a(a) = 0$ and $a_k = a$, so effectively $\tilde{I} = I$.

Moreover, taking $b = a_0$ in the result of Theorem 2.10 we have

$$I(a_0) \leq \inf(I) + K \lim_k d_\phi(a_{n_k}, a_0) = \inf(I) + d_\phi(a_0, a_0) = \inf(I),$$

and so $I(a_0) = \inf(I)$. Consequently we can write $\tilde{I}(\cdot) = I(a_0) + d_\phi(a_0, \cdot)$, being $a_0 \in D$ a point in which I attains their minimum. In addition, we can assume for the indices we work with that $\inf(I) = 0$, so we can make this approximation for any Q-normalised and ϕ-coherent index for K. In this case $\tilde{I}(\cdot) = d_\phi(a_0, \cdot)$. This issue will be explored further in Remark 2.12.

2.2.2 Extension theorems: McShane and Whitney formulas

The other approach to the problem in question that we will work on is based on using the known values of the index to construct an approximation to the unknown values, using the extension formulas that we saw in Chapter 1. Although other extension techniques are currently available that may be better for their purpose (see for example [10]), for the aim of this study it is more convenient to work with the classical McShane and Whitney methods. This is because these formulas make it easier to work with the normalisation constant and the ϕ-coherence condition of the indices they extend, as well as offering the possibility to directly compare the result of the extension with the approximations using standard indices, as we will see in the next chapter. Therefore, we will now look at the relationship between the normalisation and the ϕ-coherence condition of an index and its extension.

Proposition 2.11. Let (D_0, d, I) an index metric subspace of (D, d, I), and suppose that $I\colon D_0 \to \mathbb{R}^+$ is Q-normalized and ϕ-coherent for K. Then, if $QK \geq 1$, the Whitney extension $I^W\colon D \to \mathbb{R}^+$ is also Q-normalized and ϕ-coherent for K.

Proof. We note that Theorem 1.13 guarantees that the Whitney extension I^W is ϕ-coherent for K, so it is enough to show that I^W is Q-normalized. Take $c, d \in D$ (not necessarily in D_0).

Then, for all $a, b \in D_0$ we get

$$\begin{aligned} d_\phi(c,d) &\leq d_\phi(c,a) + d_\phi(a,b) + d_\phi(b,d) \\ &\leq d_\phi(c,a) + QI(a) + QI(b) + d_\phi(b,d) \\ &\leq QKd_\phi(c,a) + QI(a) + QI(b) + QKd_\phi(b,d) \\ &= Q(I(a) + Kd_\phi(c,a)) + Q(I(b) + Kd_\phi(b,d)), \end{aligned}$$

being latter inequality a consequence of the assumption $QK \geq 1$. Therefore,

$$\begin{aligned} d_\phi(c,d) &\leq Q \inf_{a \in D_0} \{I(a) + Kd_\phi(a,c)\} + Q \inf_{b \in D_0} \{I(b) + Kd_\phi(b,d)\} \\ &= Q(I^W(c) + I^W(d)), \end{aligned}$$

so I^W is Q-normalized. □

Remark 2.12. In case that condition $QK \geq 1$ does not hold, we can also estimate the normalization constant following the same procedure as in the previous proof. For example, if we have $0 < \alpha \leq QK$, we can ensure that the extension is (Q/α)-normalized. Nevertheless, the assumption $QK \geq 1$ is in a sense universal in index spaces. Indeed, since I represents an index, we can suppose that $I(a) \geq 0$ for all $a \in D$. Moreover, if I attains the minimum in $b \in D$ (something that happens if, for example, D is finite or compact) and this minimum is 0, then

$$\begin{aligned} I(a) = I(a) - I(b) = |I(a) - I(b)| &\leq Kd_\phi(a,b) \\ &\leq KQ(I(a) + I(b)) = KQI(a). \end{aligned}$$

Hence, if there exists $b \in D$ such that $I(b) = 0$ we can ensure that $KQ \geq 1$ as in Proposition 2.11. Moreover, even if this situation does not arise, for an index defined in a compact metric space we will have that $\inf(I) = I(b)$ for some $b \in D$, and then $I_0(a) := I(a) - I(b)$ is another positive index that preserves the same order properties as I and $I_0(b) = 0$.

For the McShane extension we can only guarantee that it preserves the coherence condition of the original index, as the normalisation constant may not be maintained. However, we can find in some cases an estimate of this constant, as we propose as follows.

Proposition 2.13. Let (D_0, d, I) an index metric subspace of (D, d, I), and suppose that $I\colon D_0 \to \mathbb{R}^+$ is Q-normalized and ϕ-coherent index for K such that $QK \geq 1$. If

$$E(D, D_0) := \sup_{c \in D} \inf_{a \in D_0} \left| \frac{I(a) + Kd_\phi(c,a)}{I(a) - Kd_\phi(c,a)} \right|$$

is finite, then the McShane extension $I^M\colon D \to \mathbb{R}$ is Q' normalized for $Q' = QE(D_0, D)$ and ϕ-coherent for K.

Proof. As in the case of Whitney's extension, Proposition 1.13 guarantees that the McShane extension I^M is ϕ-coherent for K, so it is enough to show that I^M is Q'-normalized. Bearing in mind that $QK \geq 1$, and taking $c, d \in D$, we get for all $a, b \in D_0$ that

$$\begin{aligned}
d_\phi(c,d) &\leq d_\phi(c,a) + d_\phi(a,b) + d_\phi(b,d) \\
&\leq d_\phi(c,a) + QI(a) + QI(b) + d_\phi(b,d) \\
&\leq Q\left(\tfrac{1}{Q} d_\phi(c,a) + I(a) + \tfrac{1}{Q} d_\phi(b,d) + I(b)\right) \\
&\leq Q\left(\left|\frac{I(a) + K d_\phi(c,a)}{I(a) - K d_\phi(c,a)}\right| |I(a) - K d_\phi(c,a)|\right. \\
&\quad \left. + \left|\frac{I(b) + K d_\phi(d,b)}{I(b) - K d_\phi(d,b)}\right| |I(b) - K d_\phi(d,b)|\right) \\
&\leq Q\left(\left|\frac{I(a) + K d_\phi(c,a)}{I(a) - K d_\phi(c,a)}\right| \left|I^M(c)\right| + \left|\frac{I(b) + K d_\phi(d,b)}{I(b) - K d_\phi(d,b)}\right| \left|I^M(d)\right|\right).
\end{aligned}$$

Since $a, b \in D_0$ are arbitrary, we can write

$$d_\phi(c,d) \leq Q\left(\inf_{a\in D_0}\left\{\left|\frac{I(a) + K d_\phi(c,a)}{I(a) - K d_\phi(c,a)}\right|\right\} \left|I^M(c)\right| + \inf_{b\in D_0}\left\{\left|\frac{I(b) + K d_\phi(d,b)}{I(b) - K d_\phi(d,b)}\right|\right\} \left|I^M(d)\right|\right),$$

so we conclude $d(c,d) \leq QE(D, D_0)\left(I^M(c) + I^M(d)\right)$. □

Since the McShane and Whitney extensions are minimal and maximal, respectively, and for each $\alpha \in (0,1)$ one has that $I^E := (1-\alpha)I^W + \alpha I^M$ is another extension, it will often be more convenient to use an intermediate extension like I^E. Therefore, we now ask ourselves whether we can obtain similar results to those we have just studied for these other extensions. If we are under the same hypothesis as in the previous propositions, we will again have that I^E is a ϕ-coherent index for K, which can be deduced in much the same way as in Theorem 1.13 for the Whitney case. With regard to the normalization constant, we note that

$$\begin{aligned}
I^E(a) + I^E(b) &= (1-\alpha)(I^W(a) + I^W(b)) + \alpha(I^M(a) + I^M(b)) \\
&\geq \frac{(1-\alpha)}{Q} d_\phi(a,b) + \frac{\alpha}{QE(D,D_0)} d_\phi(a,b) \\
&= \frac{(1-\alpha)E(D,D_0) + \alpha}{QE(D,D_0)} d_\phi(a,b),
\end{aligned}$$

so I^E is $\tilde{Q}$-normalized for $\tilde{Q} = \dfrac{QE(D,D_0)}{(1-\alpha)E(D,D_0) + \alpha}$.

Chapter 3

Applications

The purpose of this chapter will be to implement the theoretical tools seen in the last chapters in algorithms capable of extending a given index. More specifically, given a metric space (D, d) and an index $I \colon D_0 \subset D \to \mathbb{R}$, we will be interested in how to approximate I on $D \setminus D_0$ as well as the error and computational time of the process. We will study this from the point of view of the index extension theory developed in previous chapters by studying the difference and drawing comparisons between the classical Lipschitz index procedure and the introduced ϕ-Lipschitz concept. In addition, we will also compare these methods with another one that is common in the literature for solving this type of problems (see Chapter 6 of [11]) : neural networks. To carry out this study, we will first develop an algorithm that will gather the ϕ-Lipschitz index extension results and then analyse its performance with two concrete examples. In the first of these, we will present and contextualise a real case of practical application where indexes appear, and we will see the importance of the extension process. It is framed within the need for indicators that allow urban planners to guide their work with the aim of improving the liveability of urban space. Finally, in order to have a better comparison between the existing alternatives, we will propose concrete examples that will allow us to adequately assess the features of all of them.

3.1 Implementation

In this section we will make a methodological proposal on how to implement the theoretical content of this work to extend a given index. We start from a finite set of elements, characterised by some real variables, for which we are interested in knowing the value of a certain index I. For some of these elements the value of the index of interest is known, so we seek to extrapolate this information to estimate its value for those elements for which it is not defined. In mathematical terms, this set is a metric space D in which there is an appropriate distance defined according to the nature of the data or the problem, and there is an index defined in $D_0 \subset D$ that one wants to extend to D. To this end, we propose the following methodology.

(I) First question we need to address is whether the different nature of the variables can perturb the metric we are working with, due to the heterogeneity of their scales. We propose, to avoid this situation, to bring them all to the same scale by subtracting the minimum and dividing by the range. More precisely, suppose that $D = \{y_j\}_{j=1}^n$ and $y_j = (x_1^j, \cdots, x_m^j)$. Let $a_k := \max_j x_k^j$ and $b_k := \min_j x_k^j$ for each $k = 1, \ldots, m$. We then transform x_k^j to

$$\frac{x_k^j - b_k}{a_k - b_k}$$

for all j and k, so we will have the new variables restricted to [0,1], in the same scale.

(II) To assess the accuracy of the approximation we will make, we need a measure of the error made, which in our case will be the Root Mean Square Error (RMSE). This yields the expected absolute error and is defined as

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\tilde{I}(a_j) - I(a_j)\right)^2},$$

where $a_1, \ldots, a_n$ are the observations where we want to estimate the error, and $\tilde{I}$ is the approximation to I. However, since we do not have information on the index for the values we want to approximate, we need a strategy that allows us to estimate this error. In our case we will divide D_0, the subset of the observations with known index, into two subsets that we will call training and test. We will use the observations from the training set, consisting of seventy percent of the total observations selected randomly, to carry out the extension, while the remaining observations from the test set will be used to calculate the RMSE. Nevertheless, the randomness of this process influences the resulting error, so it may not be representative. To address this situation, we will carry out a process known as cross-validation, which consists of repeating this process several times (in our case twenty times) to compile the resulting errors and to be able to draw a more robust conclusion on accuracy. Moreover, among all the training sets that we have generated, we will take as a reference for the extension the one with the lowest resulting RMSE.

(III) To choose the best ϕ function to fit the model we will carry out an optimisation process, in which we will extract the function that minimises the error made in the test set. Ideally, we would partition our data set into three subsets, the aforementioned training and test subsets plus a validation subset, from which we would make the adjustment. However, as we do not have numerous observations, the resulting sets would not be significant for this study, so we will take the values obtained from the test set as a reference. To do so, we will consider that the linear combination of functions in Φ with positive scalars is another function in Φ. We will first choose a set of elementary functions $\{\phi_j\}_{j=1}^n \subset \Phi$

and we will discuss for which values $\lambda_1, \ldots, \lambda_n \geq 0$ the funcion $\phi := \lambda_1\phi_1 + \ldots + \lambda_n\phi_n$ ensures that the metric d_ϕ is optimal in terms of the RMSE. To do so, we will consider the particle swarm optimisation algorithm of the R library "pso". This type of algorithm, in contrast to those based on the gradient of the function, explores the entire possible set of parameters and thus avoids convergence to local minimums.

(IV) If we consider the extension using the Whitney and McShane formulas, which are maximal and minimal extensions respectively, we can ask ourselves whether we can consider an intermediate extension that minimizes the error. That is to say, for which $\alpha \in [0,1]$ the extension $I := (1-\alpha)I^W + \alpha I^M$ minimizes the error. As in the previous item, the preferred way to obtain this parameter would be from a validation set, but for the reasons already explained, we will take the test set as a reference. To do so, we will choose the value of α according to the following result.

Proposition 3.1. Let (D,d) be a finite metric space and $I\colon D_0 \subset D \to \mathbb{R}$ a ϕ-coherent index for $K > 0$. Let $S_1, S_2 \subset D_0$ such that $S_1 \cup S_2 = D_0$ and $S_1 \cap S_2 = \emptyset$. Consider

$$I^W(b) := \inf_{a \in S_1} \{I(a) + K d_\phi(a,b)\}, \quad I^M(b) := \sup_{a \in S_1} \{I(a) - K d_\phi(a,b)\}.$$

Naming $I^E_\alpha := (1-\alpha)I^W + \alpha I^M$, then $\displaystyle\min_{0 \leq \alpha \leq 1} \sum_{b \in S_2} \left(I(b) - I^E_\alpha(b)\right)^2 = \sum_{b \in S_2} \left(I(b) - I^E_{\alpha_0}(b)\right)^2$ for

$$\alpha_0 = \frac{\sum_{b \in S_2} \left(I^W(b) - I(b)\right)\left(I^W(b) - I^M(b)\right)}{\sum_{b \in S_2} \left(I^W(b) - I^M(b)\right)^2}.$$

Proof. Let $F(\alpha) := \displaystyle\sum_{b \in S_2} \left(I(b) - I^E_\alpha(b)\right)^2$ and note that

$$F(\alpha) = \sum_{b \in S_2} \left(I(b) - I^W(b) + \alpha\left(I^W(b) - I^M(b)\right)\right)^2,$$

so

$$\begin{aligned} F'(\alpha) &= 2\sum_{b \in S_2} \left(I(b) - I^W(b) + \alpha\left(I^W(b) - I^M(b)\right)\right)\left(I^W(b) - I^M(b)\right) \\ &= 2\sum_{b \in S_2} \left(I(b) - I^W(b)\right)\left(I^W(b) - I^M(b)\right) \\ &\quad + 2\alpha \sum_{b \in S_2} \left(I^W(b) - I^M(b)\right)^2. \end{aligned}$$

Solving the equation $F'(\alpha) = 0$ we obtain the α_0 we are looking for, since

$$F''(\alpha) = 2\sum_{b \in S_2} \left(I^W(b) - I^M(b)\right)^2 \geq 0.$$

$\square$

The reason why we have considered the RMSE to measure the error as opposed to other alternatives is because it ensures the derivability of the function F that we have considered, which allows us to easily study its minimum as in the previous proposition. If we had considered other types of error measures defined from the absolute value, such as the mean absolute error (MAE) or the symmetric mean absolute percentage error (SMAPE), it would have led to a less direct and more complex study to determine the optimal α due to the non-derivative nature of such definitions.

(V) In case we consider the extension by identifying the index with a standard one, we will proceed to find the element $a_0 \in D$ that minimises the index. In this case we will take $\tilde{I}(b) := K d_\phi(a_0, b)$ as an approximation of $I(b)$ because $\min(I) = 0$ after scaling.

3.2 Testing the algorithm

Having presented and studied the main issues in putting our theory into practice, we have implemented in the programming language R the algorithm that follows the points outlined above. The objective of this section is to analyse the numerical results obtained from it and to compare its performance with other existing alternatives. Specifically, we will apply these techniques in two cases: a first example that highlights the importance of index theory and the advantages of having its extension, and another that will help us to make a more detailed study of its results and comparisons. All the codes that have been used can be found in the Appendix 3.2.2.

3.2.1 AARP livability index

In 2018, 55 percent of the world's population lived in urban areas, and by 2050 this ratio is expected to rise to 68 percent, according to [12]. This context of rapid urbanisation explains the growing interest in studying and measuring concepts such as the quality of life or liveability of cities, as can be seen, for example, in the different indices summarized in [13]. The purpose of these indicators is twofold: firstly, to define what is meant by liveability and to identify the parameters that describe it, and secondly, to have information on which cities or neighbourhoods have better living conditions. With this information, urban planners can better understand the areas in which to act, or public administrations can identify places with poorer living conditions to develop and in which to invest. However, this task can be difficult to accomplish due to the numerous and hard-to-estimate factors involved. For example, one of the best-known liveability indices, the Global Liveability Index compiled by The Economist, lists thirty indicators in five different categories, some of which are in turn built upon others. Moreover, some of them are subjective and difficult to estimate, such as the discomfort of climate for travellers or level of corruption. In this section, we propose to use the index extension theory developed in the previous chapter to approximate liveability using only alternative mobility indicators to the private car. The idea behind this is to dispense with subjective or complex to estimate social indicators, and to focus on these easily estimable ones based on existing infrastructure, and the connectivity of the urban pattern beyond car dependency and its associated problems.

Let us explain the databases we will consider. Walk Score®[1] is a website that scores, from 0 to 100, the walkability performance of 123 cities in the United States and Canada according to a series of parameters, such as intersection density, block length, access to amenities in less than a 5-minute walk, etc. Also, it provides scores on cities' transport and cycling performance in the same way. Our objective will be to use these three indicators to approximate the AARP liveability index [2]. This index assesses 61 different indicators in seven categories (housing, neighbourhood, transport, environment, health, engagement and opportunity) to evaluate the liveability of US cities, such as housing costs, crime rate, air quality or income inequality. The final score is a number between 0 and 100, 50 being and average score and higher values corresponding to above-average performance, and vice versa. In mathematical terms our metric space is $D \subset [0, 100]^3$, where each element $(x, y, z) \in D$ represents a city with walk, transport and bike scores x, y and z respectively, equipped with the canonical metric d of $\mathbb{R}^3$. For 101 US cities we have defined the index of interest, which we will call I, and our goal is to be able to define it for 22 Canadian cities as well. Table 3.1 shows an example of our data.

We have implemented in Algorithm 1 the extension of our index, following the considerations set out in the previous section. In particular, to determine the ϕ function, we consider two linear combination of funcions of Φ as follows:

$$\phi(x) = p_1 x + p_2 \log(1 + x) + p_3 \arctan(x) + p_4 \frac{x}{1 + x}, \quad p_1, \dots, p_4 \geq 0,$$

and

$$\psi(x) = p_1 \sqrt{x} + p_2 \log\left(1 + \sqrt{x}\right) + p_3 \arctan\left(\sqrt{x}\right) + p_4 \frac{\sqrt{x}}{1 + \sqrt{x}}, \quad p_1, \dots, p_4 \geq 0.$$

Table 3.2 shows the comparative performance of the two methods (identification with a standard index and McShane-Whitney extensions) and their differences with of our technique (introduction of a function of Φ). As we can see, the identification of our index with a standard one gives a bad performance due to a significantly high estimated mean error, so we cannot consider it. As for the extension formulas, our technique maintains the standard deviation of the error while reducing the expected RMSE, although not significantly. Furthermore, the computation time is higher, mainly due to the optimisation process carried out. On the other hand, Table 3.3 shows the results of considering in this case a linear combination as ψ, together with the same results of the classical technique and the neural net. In this case, there is a more significant improvement in performance compared to the original technique. However, the neural network has been shown to be the most efficient in terms of prediction. Finally, in

[1] https://www.walkscore.com
[2] https://livabilityindex.aarp.org/scoring

Table 3.4 we present the predictions resulting from each method, providing a ranking according to them. Since we do not have benchmark values for our studied index when assessing the resulting rankings, we will compare it with other existing indices. The Mercer quality of living city ranking[3] classifies Canadian cities in order as follows: Vancouver, Toronto, Ottawa, Montreal and Calgary. As can be seen, this ranking is reasonably consistent with the ones we have offered, with the city of Calgary being the only one that differs. Finally, we note that the positive results obtained confirm that the model we have proposed is consistent, that is, it is possible to characterise liveability based on parameters related to the mobility of pedestrians, cyclists and public transport.

3.2.2 Further analysis

In order to check that the conclusions we have reached above are generally valid, and to provide further examples to validate the standard index identification method, we will now present some datasets to study this. We will first explain how we have generated our datasets.

We start with four random variables with a normal distribution, mean 0 and standard deviation 1, from which we have obtained 100 random values and which we have multiplied by 10 and rounded. We have called them x, y, z and t. Our metric space is formed by each of the vectors of the form (x, y, z, t), equipped with the usual metric of $\mathbb{R}^4$. The index that we have made corresponding to each of these elements is obtained by rounding the distance between them and a preset element, plus the sine of this distance. In table 3.5 we have provided examples of the observations we have generated and their indices. The purpose of doing so is to obtain an index similar to a standard one (the coherence constant obtained after scaling is $K = 1.378$ and the normalization constant $Q = 0.994$, approximately) and in this case to be able to obtain a better comparison. We will also do the same but with 200 observations, so that we can study the performance of the different alternatives as we increase the size of our problem. Tables 3.6 and 3.7 show the results obtained.

Note that the introduction of a function $\phi \in \Phi$ reduces the expected error of the extensions, this fact being more significant in the standard index method, where the standard deviation is also considerably reduced. As a counterpart, we see how the computation time increases again in order of 3. Moreover, unlike the previous case, we can see now how the neural network no longer has the best performance, the most significant fact being the considerable standard deviation of the results it obtains. Finally, we note that doubling the number of observations also approximately doubles the execution time of the classical Lipschitz methods, while for the new proposed methods it is multiplied by a factor of 3. Moreover, we can see that the introduction of one more variable compared to the previous problem has hardly changed the computational times.

[3] https://mobilityexchange.mercer.com/Insights/quality-of-living-rankings

City	Walk Score	Transit Score	Bike Score	I
New York	88	88.6	69.3	63
Los Angeles	68.6	52.9	58.7	49
Chicago	77.2	65	72.2	57
Toronto	61	78.2	61	?
Houston	47.5	36.2	48.6	48
Montreal	65.4	67	72.6	?

Table 3.1: Examples of scores and index for some cities.

Function ϕ	Standard		McShane-Whitney	
	Lipschitz	ϕ-Lipschitz	Lipschitz	ϕ-Lipschitz
Mean RMSE	138.03	84.59	6.621	6.547
Median RMSE	142.06	85.28	6.704	6.660
Standard deviation	27.909	12.720	0.653	0.614
Seconds per iteration	1.548×10^{-4}	2.396×10^{-1}	5.875×10^{-4}	4.579×10^{-1}

Table 3.2: Comparison of method performance for function ϕ.

Function ψ	Standard		McShane-Whitney		Neural net
	Lipschitz	ψ-Lipschitz	Lipschitz	ψ-Lipschitz	
Mean RMSE	138.03	20.17	6.621	5.777	3.943
Median RMSE	142.06	19.65	6.704	5.750	3.919
Standard deviation	27.909	3.001	0.653	0.683	0.512
Seconds per iteration	1.548×10^{-4}	2.421×10^{-1}	5.875×10^{-4}	4.245×10^{-1}	1.230×10^{-1}

Table 3.3: Comparison of method performance for function ψ.

Ranking	Classic Lipschitz		ϕ-Lipschitz		Neural net	
	City	Index	City	Index	City	Index
1	Vancouver	65	Vancouver	62	Toronto	62
2	Toronto	65	Longueuil	62	Vancouver	62
3	Longueuil	62	Montreal	60	Burnaby	60
4	Ottawa	61	Winnipeg	58	Montreal	60
5	Montreal	61	Toronto	58	Ottawa	58
6	Winnipeg	61	Saskatoon	56	Winnipeg	58
7	Hamilton	58	Ottawa	56	Longueuil	57
8	Saskatoon	57	Hamilton	55	Mississauga	57
9	Surrey	56	Kitchener	54	Brampton	57
10	Laval	56	Surrey	54	Laval	57
11	Brampton	56	London	54	Markham	56
12	Quebec	55	Brampton	54	Calgary	56
13	Kitchener	55	Mississauga	53	Surrey	56
14	London	54	Laval	53	Kitchener	56
15	Calgary	54	Edmonton	53	Hamilton	55
16	Mississauga	53	Quebec	53	Vaughan	55
17	Gatineau	52	Windsor	53	Windsor	55
18	Windsor	51	Gatineau	52	Quebec	53
19	Edmonton	50	Calgary	52	London	52
20	Burnaby	50	Vaughan	51	Edmonton	52
21	Vaughan	50	Burnaby	50	Gatineau	52
22	Markham	49	Markham	50	Saskatoon	50

Table 3.4: Ranking of Canadian cities by predicted AARP liveability index.

x	y	z	t	Index
6	-7	22	-7	35
-2	3	13	-8	27
16	-2	-3	-9	28
1	-3	5	-11	25
1	-10	-4	-4	22
17	0	-5	3	26

Table 3.5: Example of some values of our dataset and its indices.

$n = 100$	Standard		McShane-Whitney		Neural net
	Lipschitz	ϕ-Lipschitz	Lipschitz	ϕ-Lipschitz	
Mean RMSE	5.462	3.344	8.332	8.266	7.6779
Median RMSE	5.470	3.335	8.382	8.322	8.1556
Standard deviation	0.433	0.242	1.189	1.174	3.631
Seconds per iteration	1.261×10^{-4}	2.217×10^{-1}	4.599×10^{-4}	4.338×10^{-1}	5.757×10^{-2}

Table 3.6: Comparison of method performance for $n = 100$.

$n = 200$	Standard		McShane-Whitney		Neural net
	Lipschitz	ϕ-Lipschitz	Lipschitz	ϕ-Lipschitz	
Mean RMSE	5.106	2.957	8.506	8.494	4.5950
Median RMSE	5.071	2.975	8.4340	8.415	3.9713
Standard deviation	0.201	0.115	0.663	0.668	3.835
Seconds per iteration	3.247×10^{-4}	7.901×10^{-1}	8.740×10^{-4}	1.411×10^{0}	7.151×10^{-2}

Table 3.7: Comparison of method performance for $n = 200$.

Conclusions

We will conclude this work by outlining the main points that have been made, as well as the conclusions of the results we have presented. In addition, we will indicate what we believe are the most interesting directions that can be taken in order to extend this work and search for new results.

In our work we have started from the Lipschitz condition, and we have made a proposal for generalisation ensuring that this new class of functions still admits extension formulas like the classical ones. As a result, we have found a way to redefine the distance in the metric space that allows us to improve the extension results. We have taken this study to the field of index spaces, presenting this theoretical framework and also contextualising possible applications and the need for index extension. The results we have seen allow us to conclude that the introduction of the ϕ-Lipschitz concept allows for a general improvement of the extensions. However, the degree to which they do so will largely depend on the nature of the index and the ϕ-function under consideration.

On the other hand, we consider that in the future it would be interesting to study possible relations between ϕ and the metric space, in order to get a convenient metric d_ϕ to make the extension. We have done this by considering linear combinations of some of the functions of Φ, optimising the value of the scalars. Although this can be improved by considering larger sets of elementary functions, for example compositions of them, we consider a problem of interest to analyse whether a description of ϕ can be obtained a priori. Furthermore, although in our basic metric spaces we have always considered the Euclidean metric, in some problems other type of distances are used, which may include weights in their components depending on the importance of the variables. It remains an open problem to study the behaviour of our method with other metrics, especially how the weights are affected when composing with a ϕ function.

Algorithms

Algorithm 1: Extension of the AARP livability index.

```
################### READING DATA ####################

#Libraries
library(dplyr)
library(Rfast)
library(neuralnet)
library(pso)

#Work directory
setwd("C:/Users/gonza/OneDrive/Documentos/lipschitz")

#Data frame
df <- read.csv("~/lipschitz/ciudades.csv", dec=",")
n <- nrow(df)

################### DATA PROCESSING ####################

#Scaling
maxs <- as.vector(apply(df[2:5], 2, max))
mins <- as.vector(apply(df[2:5], 2, min))
scaled <- data.frame(scale(df[,2:5], mins, maxs-mins))
colnames(scaled) <- c("walk","transit","bike","liveability")

#Matrix of distances
D <- Dist(scaled[,1:3])

#Index distances
Dind <- Dist(scaled[,4])

#Optim parameters
p0 <- c(1,0,0,0)
low <- c(1e-16,1e-16,1e-16,1e-16)
```

```
up <- c(1,1,1,1)

################### EXTENSION ####################

N <- 20
e.standard <- NULL
e.phistandard <- NULL
e.mw <- NULL
e.phimw <- NULL
e.nn <- NULL
t.standard <- 0
t.phistandard <- 0
t.mw <- 0
t.phimw <- 0
t.nn <- 0

for (iter in 1:N) {

  #Training and test
  train <- sort(sample(1:n,round(0.7*n)))
  test <- setdiff(1:n,train)

  #Standard Lipschitz
  t0 <- Sys.time()
  L <- max(Dind[train,]/D[train,], na.rm=TRUE)
  a0 <- which.min(scaled[train,4])
  standard <- L*D[a0,test]
  standard <- standard*(maxs[4]-mins[4])+mins[4]
  e.standard[iter] <- sqrt(sum((standard-df$liveability[test])^2)/length(test))
  t1 <- Sys.time()
  t.standard <- t.standard + t1-t0

  #Standard phi-Lipschitz
  t0 <- Sys.time()
  a0 <- which.min(scaled[train,4])
  J <- function(p) {

    phi <- function(x) {
      x <- sqrt(x)
      func <- p[1]*x + p[2]*log(1+x) + p[3]*atan(x) + p[4]*x/(1+x)
      return(func)
    }

    K <- max(Dind[train,]/phi(D[train,]), na.rm = TRUE)
```

```r
  standard <- K*phi(D[a0,test])
  standard <- standard*(maxs[4]-mins[4])+mins[4]
  error <- sqrt(sum((standard-df$liveability[test])^2)/length(test))
  return(error)
}
optimum <- psoptim(p0, J, lower=low, upper=up,
                   control = list(maxit = 100, maxf = 200))
e.phistandard[iter] <- optimum$value
t1 <- Sys.time()
t.phistandard <- t.phistandard + t1-t0

#McShane-Whitney Lipschitz
t0 <- Sys.time()
L <- max(Dind[train,]/D[train,], na.rm = TRUE)
index <- scaled$liveability[train]
Whitney <- apply(index + L*D[test,train], 1, min)
McShane <- apply(index - L*D[test,train], 1, max)
dif <- Whitney - McShane
alpha <- sum(dif*(Whitney-scaled$liveability[test]))/sum(dif^2)
I <- (1-alpha)*Whitney + alpha*McShane
I <- I*(maxs[4]-mins[4])+mins[4]
e.mw[iter] <- sqrt(sum((I-df$liveability[test])^2)/length(test))
t1 <- Sys.time()
t.mw <- t.mw + t1-t0

if (e.mw[iter] == min(e.mw)) {
  Lip <- L
  Alpha <- alpha
}

#McShane-Whitney phi-Lipschitz
t0 <- Sys.time()
J <- function(p) {

  phi <- function(x) {
    x <- sqrt(x)
    func <- p[1]*x + p[2]*log(1+x) + p[3]*atan(x) + p[4]*x/(1+x)
    return(func)
  }

  K <- max(Dind[train,]/phi(D[train,]), na.rm = TRUE)
  index <- scaled$liveability[train]
  Whitney <- apply(index + K*phi(D[test,train]), 1, min)
  McShane <- apply(index - K*phi(D[test,train]), 1, max)
```

```
    dif <- Whitney - McShane
    alpha <- sum(dif*(Whitney-scaled$liveability[test]))/sum(dif^2)
    I <- (1-alpha)*Whitney + alpha*McShane
    I <- I*(maxs[4]-mins[4])+mins[4]
    error <- sqrt(sum((I-df$liveability[test])^2)/length(test))
    return(error)
  }
  optimum <- psoptim(p0, J, lower=low, upper=up,
                     control = list(maxit = 100, maxf = 200))
  e.phimw[iter] <- optimum$value
  t1 <- Sys.time()
  t.phimw <- t.phimw + t1-t0

  if (e.phimw[iter] == min(e.phimw)) {
    Train <- train
    Test <- test
    par <- optimum$par
  }

  #Neural Net
  t0 <- Sys.time()
  nn <- neuralnet(liveability ~ walk + transit + bike, scaled[train,])
  pr.nn <- compute(nn, scaled[test,1:3])
  pr.nn <- pr.nn$net.result*(maxs[4]-mins[4])+mins[4]
  e.nn[iter] <- sqrt(sum((pr.nn-df$liveability[test])^2)/length(test))
  t1 <- Sys.time()
  t.nn <- t.nn + t1-t0

  if (e.nn[iter] == min(e.nn)) {
    NN <- nn
  }

}

#Summary
summary(e.standard)
summary(e.phistandard)
summary(e.mw)
summary(e.phimw)
summary(e.nn)

#Times
t.standard/N
t.phistandard/N
```

```
t.mw/N
t.phimw/N
t.nn/N

################### CANADA RANKING ####################

#Canada data frame
df.can <- read.csv("~/lipschitz/ciudadescan.csv", dec=",")
m <- nrow(df.can)

#Predictions
index <- scaled$liveability[Train]

#McShane-Whitney Lipschitz
MW <- function(city) {
  city <- (city-mins[1:3])/(maxs[1:3]-mins[1:3])
  distance <- sqrt(apply((scaled[Train,1:3]-city)^2,1,sum))
  Whitney <- min(index + Lip*distance)
  McShane <- max(index - Lip*distance)
  index = (1-Alpha)*Whitney + Alpha*McShane
  index <- index*(maxs[4]-mins[4])+mins[4]

  return(index)
}

#McShane-Whitney phi-Lipschitz
phi <- function(x) {
  func <- par[1]*x + par[2]*log(1+x) + par[3]*atan(x) + par[4]*x/(1+x)
}

K <- max(Dind[Train,]/phi(D[Train,]), na.rm = TRUE)
Whitney <- apply(index + K*phi(D[Test,Train]), 1, min)
McShane <- apply(index - K*phi(D[Test,Train]), 1, max)
dif <- Whitney - McShane
Alpha.phi <- sum(dif*(Whitney-scaled$liveability[Test]))/sum(dif^2)

phiMW <- function(city) {
  city <- (city-mins[1:3])/(maxs[1:3]-mins[1:3])
  distance <- sqrt(apply((scaled[Train,1:3]-city)^2,1,sum))
  Whitney <- min(index + K*phi(distance))
  McShane <- max(index - K*phi(distance))
  index = (1-Alpha.phi)*Whitney + Alpha.phi*McShane
  index <- index*(maxs[4]-mins[4])+mins[4]
```

```r
  return(index)
}

#Rankings
ranking <- data.frame(df.can$city)
for (i in 1:m){
  ranking[i,2] <- MW(as.numeric(df.can[i,2:4]))
}

ranking[,3] <- df.can$city
for (i in 1:m){
  ranking[i,4] <- phiMW(as.numeric(df.can[i,2:4]))
}

ranking[,5] <- df.can$city
pred <- compute(NN,(df.can[,2:4]-mins[1:3])/(maxs[1:3]-mins[1:3]))
pred <- pred$net.result
ranking[,6] <- pred*(maxs[4]-mins[4])+mins[4]
colnames(ranking) <- c("City","Index","City","Index","City","Index")
write.csv(ranking, "C:/Users/gonza/OneDrive/Documentos/lipschitz/ranking.csv",
          row.names=FALSE)
```

Algorithm 2: Extension of our custom index.

```
################### READING DATA ####################

#Libraries
library(dplyr)
library(Rfast)
library(neuralnet)
library(pso)

#Work directory
setwd("C:/Users/gonza/OneDrive/Documentos/lipschitz")

#Data frame
df <- read.csv("~/lipschitz/datos100.csv", dec=",")
n <- nrow(df)

################### DATA PROCESSIN ####################

#Scaling
maxs <- apply(df, 2, max)
mins <- apply(df, 2, min)
scaled <- as.data.frame(scale(df, center = mins, scale = maxs - mins))

#Minimum element
a0 <- which.min(scaled$Index)

#Matrix of distances
D <- Dist(scaled[,1:4])

#Index distances
Dind <- Dist(scaled[,5])

#Optim parameters
p0 <- c(1,0,0,0)
low <- c(1e-16,1e-16,1e-16,1e-16)
up <- c(1,1,1,1)

#################### COMPARATIONS ###################

N <- 20

RMSE_stand <- NULL
```

```
RMSE_stand_phi <- NULL
RMSE_MW <- NULL
RMSE_MW_phi <- NULL
RMSE_neuralnet <- NULL

stand_time <- 0
MW_time <- 0
stand_phi_time <- 0
MW_phi_time <- 0
neuralnet_time <- 0

for (iter in 1:N) {

  #Train and test
  train <- sort(sample(1:n,round(0.7*n)))
  test <- setdiff(1:n,train)

  ### Standard Lipschitz
  t0 <- Sys.time()
  K <- max(Dind[train,]/D[train,], na.rm = TRUE)
  standard <- K*D[a0,test]
  standard <- standard*(maxs[5]-mins[5])+mins[5]
  RMSE_stand[iter] <- sqrt(sum((standard-df$Index[test])^2)/length(test))
  t1<- Sys.time()
  stand_time <- stand_time + t1-t0

  ###McShane-Whitney Lipschitz
  t0 <- Sys.time()
  index <- scaled$Index[train]
  Whitney <- apply(index + K*D[test,train], 1, min)
  McShane <- apply(index - K*D[test,train], 1, max)
  dif <- Whitney - McShane
  alpha <- sum(dif*(Whitney-scaled$Index[test]))/sum(dif^2)
  I <- (1-alpha)*Whitney + alpha*McShane
  I <- I*(maxs[5]-mins[5])+mins[5]
  RMSE_MW[iter] <- sqrt(sum((I-df$Index[test])^2)/length(test))
  t1<- Sys.time()
  MW_time <- MW_time + t1-t0

  ### Standard phi-Lipschitz
  t0 <- Sys.time()
  J <- function(p) {

    phi <- function(x) {
```

```
    u <-p[1]*x + p[2]*log(1+x) + p[3]*atan(x) + p[4]*x/(1+x)
    return(u)
  }

  K <- max(Dind[train,]/phi(D[train,]), na.rm = TRUE)

  standard <- K*phi(D[a0,test])
  standard <- standard*(maxs[5]-mins[5])+mins[5]

  rmse <- sqrt(sum((standard-df$Index[test])^2)/length(test))

  return(rmse)
}

optimum<-psoptim(p0, J, lower = low, upper = up,
               control = list(maxit = 100, maxf = 200))
RMSE_stand_phi[iter] <- optimum$value
t1<- Sys.time()
stand_phi_time <- stand_phi_time + t1-t0

###McShane-Whitney phi-Lipschitz
t0 <- Sys.time()
J <- function(p) {

  phi <- function(x) {
    u <-p[1]*x + p[2]*log(1+x) + p[3]*atan(x) + p[4]*x/(1+x)
    return(u)
  }

  K <- max(Dind[train,]/phi(D[train,]), na.rm = TRUE)
  index <- scaled$Index[train]
  Whitney <- apply(index + K*phi(D[test,train]), 1, min)
  McShane <- apply(index - K*phi(D[test,train]), 1, max)
  dif <- Whitney - McShane
  alpha <- sum(dif*(Whitney-scaled$Index[test]))/sum(dif^2)
  I <- (1-alpha)*Whitney + alpha*McShane
  I <- I*(maxs[5]-mins[5])+mins[5]
  rmse <- sqrt(sum((I-df$Index[test])^2)/length(test))

  return(rmse)
}

optimum<-psoptim(p0, J, lower = low, upper = up,
                control = list(maxit = 100, maxf = 200))
```

```
  RMSE_MW_phi[iter] <- optimum$value
  t1<- Sys.time()
  MW_phi_time <- MW_phi_time + t1-t0

  ###Neural Net
  t0 <- Sys.time()
  nn <- neuralnet(Index ~ x + y + z + t, data = scaled[train,])
  pr.nn <- compute(nn, scaled[test,1:4])
  pr.nn <- pr.nn$net.result*(maxs[5]-mins[5])+mins[5]
  RMSE_neuralnet[iter] <- sqrt((sum(pr.nn-df$Index[test])^2)/length(test))
  t1<- Sys.time()
  neuralnet_time <- neuralnet_time + t1-t0
}

#Summary
summary(RMSE_stand)
summary(RMSE_MW)
summary(RMSE_stand_phi)
summary(RMSE_MW_phi)
summary(RMSE_neuralnet)

#Times
stand_time/N
MW_time/N
stand_phi_time/N
MW_phi_time/N
neuralnet_time/N
```

Bibliography

[1] Haim Brézis. *Functional analysis, Sobolev spaces and partial differential equations*, volume 2. Springer, 2011. URL https://doi.org/10.1007/978-0-387-70914-7.

[2] Pierre Hansen and Brigitte Jaumard. *Lipschitz Optimization*, pages 407–493. Springer, 1995. URL https://doi.org/10.1007/978-1-4615-2025-2_9.

[3] J.M. Calabuig, H. Falciani, and E.A. Sánchez-Pérez. Dreaming machine learning: Lipschitz extensions for reinforcement learning on financial markets. *Neurocomputing*, 398:172–184, 2020. URL https://doi.org/10.1016/j.neucom.2020.02.052.

[4] Jonathan Rosenberg. Applications of analysis on lipschitz manifolds. *Proc. Miniconferences on Harmonic Analysis and Operator Algebras (Canberra, t987), Proc. Centre for Math. Analysis*, 16:269–283, 1988.

[5] Gad Levanon. Evaluating and comparing leading and coincident economic indicators. *Business Economics*, 45(1):16–27, 2010. URL https://doi.org/10.1057/be.2009.29.

[6] Antonia Ferrer-Sapena, Ezgi Erdoğan, Eduardo Jiménez-Fernández, Enrique Alfonso Sánchez-Pérez, and Fernanda Peset. Self-defined information indices: application to the case of university rankings. *Scientometrics*, 124:2443–2456, 2020. URL https://doi.org/10.1007/s11192-020-03575-6.

[7] Elena Deza, Michel Marie Deza, Michel Marie Deza, and Elena Deza. *Encyclopedia of distances*. Springer, 2009. URL https://doi.org/10.1007/978-3-642-00234-2.

[8] Ştefan Cobzaş, Radu Miculescu, and Adriana Nicolae. *Lipschitz Functions*. Springer International Publishing, 2019. URL https://doi.org/10.1007/978-3-030-16489-8.

[9] Ezgi Erdoğan, Antonia Ferrer-Sapena, Eduardo Jiménez-Fernández, and Enrique A. Sánchez-Pérez. Index spaces and standard indices in metric modelling. *Nonlinear Analysis: Modelling and Control*, 27(4):803–822, 2022. URL https://www.doi.org/10.15388/namc.2022.27.27493.

[10] Petri Juutinen. Absolutely minimizing lipschitz extensions on a metric space. *Annales Academiae Scientiarum Fennicae. Mathematica*, 27(1):57–67, 2002. URL http://eudml.org/doc/124719.

[11] Miroslav Kubat. *An introduction to machine learning*, volume 2. Springer, 2017. URL https://doi.org/10.1007/978-3-319-63913-0.

[12] Hannah Ritchie and Max Roser. Urbanization. *Our World in Data*, 2018. URL https://ourworldindata.org/urbanization.

[13] Arpan Paul and Joy Sen. A critical review of liveability approaches and their dimensions. *Geoforum*, 117:90–92, 2020. URL https://doi.org/10.1016/j.geoforum.2020.09.008.