

# On the use of restricted dissimilarity and dissimilarity-like functions for defining penalty functions

G. Beliakov<sup>1</sup> H. Bustince, J. Fernandez<sup>2</sup> R. Mesiar<sup>3</sup> A. Pradera<sup>4</sup>

<sup>1</sup>Deakin University, Melbourne, Australia, gleb@deakin.edu.au

<sup>2</sup>Universidad Publica de Navarra, Pamplona, Spain, bustince,fcojavier.fernandez@unavarra.es

<sup>3</sup>Slovak Institute of Technology, Bratislava, Slovakia

<sup>3</sup>UTIA AV CR Prague, P.O.Box 18, 18208 Prague, Czech Republic, mesiar@math.sk

<sup>4</sup>Universidad Rey Juan Carlos, Madrid, Spain, ana.pradera@urjc.es

## Abstract

In this work we study the relation between restricted dissimilarity functions-and, more generally, dissimilarity-like functions- and penalty functions and the possibility of building the latter using the former. Several results on convexity and quasi-convexity are also considered.

**Keywords:** Restricted dissimilarity functions, penalty functions, aggregation functions, convexity, maxitivity.

## 1. Introduction

Penalty functions provide a very useful tool to obtain an output which is the most similar one to a given set of inputs ([6, 7, 11]). In this sense, they are very useful in any application in which several inputs have to be merged into a single output containing the relevant information provided by the inputs. This usefulness has been made clear in many different applications, such as image processing (for the fusion of different images) [1] or decision making (to evaluate each of the alternatives taking into account the different criteria)[3]. Please note that the notion of penalty function here is not the same as the commonly used in optimization.

A key question is how these penalty functions can be built. In [4] it was established that a possible manner is to consider a weighted mean of the so-called faithful dissimilarity functions, which can be built from some continuous strictly monotone function (scaling function) and a convex function (shape function) which has a unique minimum at the origin.

But, since a weighted mean is a particular instance of aggregation function, this approach raises the question of whether this construction can be generalized by using other aggregation functions or even other kinds of dissimilarity-like functions.

In this work we carry on a study on restricted dissimilarity functions, their relation with convexity and quasi-convexity and their use to build penalty functions. Finally we also consider the relaxation

of the conditions required to restricted dissimilarity functions.

The structure of the contribution is as follows. We start with some preliminaries including some properties of penalty functions and penalty-based aggregation functions. In Section 3 we deal with restricted dissimilarity functions and we analyze their relation with metrics and convexity. Section 4 and 5 are devoted to the construction of penalty functions by means of aggregation functions and dissimilarity-like functions. We finish with some conclusions.

## 2. Preliminaries

**Definition 1** [2, 9] *An aggregation function is a mapping  $M : [0, 1]^n \rightarrow [0, 1]$  which is monotone increasing and such that*

$$M(0, \dots, 0) = 0 \text{ and } M(1, \dots, 1) = 1$$

Well-known examples of aggregation functions are t-norms or t-conorms. For us, the following definition is of particular interest.

**Definition 2** *A mean or averaging aggregation function is an aggregation function  $M : [0, 1]^n \rightarrow [0, 1]$  such that*

$$\min(x_1, \dots, x_n) \leq M(x_1, \dots, x_n) \leq \max(x_1, \dots, x_n)$$

for every  $(x_1, \dots, x_n) \in [0, 1]^n$ .

Note that means and idempotent aggregation functions (i.e.,  $M(x, \dots, x) = x$  for every  $x \in [0, 1]$ ) are the same. As examples of averaging aggregation functions we have the arithmetic mean:

$$M(x_1, \dots, x_n) = \frac{x_1 + \dots + x_n}{n}$$

or, more generally, the so-called Kolmogorov means [10, 12]

$$M(x_1, \dots, x_n) = f^{-1}(w_1 f(x_1) + \dots + w_n f(x_n))$$

where  $(w_1, \dots, w_n) \in [0, 1]^n$  is a weighting vector (i.e.,  $w_1 + \dots + w_n = 1$ ) and  $f : [0, 1] \rightarrow [-\infty, \infty]$  is a strictly monotone mapping. We will also deal with

fuzzy negations, which are non-increasing functions  $N : [0, 1] \rightarrow [0, 1]$  such that  $N(0) = 1$  and  $N(1) = 0$  that are said to be *strict negations* when they are strictly decreasing.

We recall now the concepts of convexity and quasi-convexity, which are crucial for the present work. Let  $D$  denote an interval.

**Definition 3** Let  $g : D \subseteq (-\infty, \infty) \rightarrow [-\infty, \infty]$  be a mapping.

1.  $g$  is convex if the inequality

$$g(\lambda x + (1 - \lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y)$$

holds for every  $x, y \in D$  and  $\lambda \in [0, 1]$ .

2.  $g$  is quasi-convex if the inequality

$$g(\lambda x + (1 - \lambda)y) \leq \max(g(x), g(y))$$

holds for every  $x, y \in D$  and  $\lambda \in [0, 1]$ .

Note that every convex function is in particular quasi-convex, whereas the reciprocal is not true. On the other hand, regarding quasi-convex functions, we have the following important result.

**Theorem 1** Let  $g : D \subseteq (-\infty, \infty) \rightarrow [-\infty, \infty]$  be a quasi-convex function. Then, the set of minimizers of  $g$  is a non-empty subinterval of  $D$ .

A penalty based function provides a way for obtaining an output which is as similar as possible to the considered inputs (in the sense of the penalty function that is chosen). We start recalling the concept of penalty function.

**Definition 4** [6, 7, 11]

A penalty function is a mapping  $P : [0, 1]^{n+1} \rightarrow [0, \infty]$  such that:

1.  $P(x_1, \dots, x_n, y) = 0$  if  $x_i = y$  for every  $i = 1, \dots, n$ ;
2.  $P(\cdot, y)$  is a quasi-convex function in  $y$ . That is, for any fixed  $(x_1, \dots, x_n) \in [0, 1]^n$  the function  $p(y) = P(x_1, \dots, x_n, y)$  is quasi-convex.

**Definition 5** Let  $P$  be a penalty function. The penalty based function (or function based on the penalty function  $P$ ) is the mapping  $f : [0, 1]^n \rightarrow [0, 1]$  defined by

$$f(x_1, \dots, x_n) = \arg \min_y P(x_1, \dots, x_n, y)$$

if  $P$  attains its minimum at a single point, or  $f(x_1, \dots, x_n) = \frac{a+b}{2}$  if  $a$  and  $b$  are the boundary points of the set of minimizers of  $P$ .

Note that in general penalty based functions need not be aggregation functions since monotonicity can not be assured. However, the following result holds.

**Theorem 2** [6] Any averaging aggregation function can be represented as a penalty based aggregation function.

### 3. Restricted dissimilarity functions, convexity and metrics

The concept of restricted dissimilarity function was introduced in [4]

**Definition 6** A restricted dissimilarity function is a mapping  $d_R : [0, 1]^2 \rightarrow [0, 1]$  such that

1.  $d_R(x, y) = d_R(y, x)$  for every  $x, y \in [0, 1]$ ;
2.  $d_R(x, y) = 1$  if and only if  $\{x, y\} = \{0, 1\}$ ;
3.  $d_R(x, y) = 0$  if and only if  $x = y$ ;
4. for any  $x, y, z \in [0, 1]$  such that  $x \leq y \leq z$  it holds that  $d_R(x, y) \leq d_R(x, z)$  and  $d_R(y, z) \leq d_R(x, z)$ .

**Example 1** [3]

1. Let  $c \in [0, 1[$ . Then the mapping  $d_R : [0, 1]^2 \rightarrow [0, 1]$  defined by:

$$d_R(x, y) = \begin{cases} 1 & \text{if } \{x, y\} = \{0, 1\}; \\ 0 & \text{if } x = y; \\ c & \text{in other case,} \end{cases}$$

is an example of restricted dissimilarity function which is not even continuous.

2. The mapping  $d_R(x, y) = \varphi_1(|\varphi_2(x) - \varphi_2(y)|)$ , where  $\varphi_1, \varphi_2$  are two automorphisms of the unit interval (i.e., two increasing bijections on  $[0, 1]$ ) is a restricted dissimilarity function.

Restricted dissimilarity function can be related to well-known functions. For instance:

**Theorem 3** Let  $d : [0, 1]^2 \rightarrow [0, 1]$  be a function. The following statements are equivalent.

- (i)  $N : [0, 1] \rightarrow [0, 1]$  is strict negation and  $d(x, y) = |N(x) - N(y)|$ ;
- (ii)  $d$  is a restricted dissimilarity function and  $d(x, y) = |d(x, 1) - d(y, 1)|$ ;
- (iii)  $d$  is a restricted dissimilarity function and for all  $x, y, z \in [0, 1]$  with  $x \geq y \geq z$ , it holds  $d(x, y) + d(y, z) = d(x, z)$  and  $d(1, x)$  is strictly monotone.

*Proof.* (i)  $\Rightarrow$  (ii): Suppose that  $d(x, y) = |N(x) - N(y)|$ . Then, symmetry of  $d$  is obvious. Moreover,  $d(x, y) = 0$  if and only if  $N(x) = N(y)$ , and since  $N$  is strict, this can happen if and only if  $x = y$ . On the other hand,  $d(x, y) = 1$  if and only if  $\{N(x), N(y)\} = \{0, 1\}$ , and once again from the strictness of  $N$ , this is equivalent to  $\{x, y\} = \{0, 1\}$ . Finally, if  $x \leq y \leq z$ , then  $d(x, y) = |N(x) - N(y)| = N(x) - N(y) \leq N(x) - N(z) = |N(x) - N(z)|$ . The case  $d(y, z) \leq d(x, z)$  is analogous. Observe that if  $x < y < z$ , then  $d(x, y) < d(x, z)$  and  $d(y, z) < d(x, z)$ , due again to the strictness of  $N$ . Finally,  $N(x) = |N(x) - N(1)| = d(x, 1)$ , so  $d(x, y) = |d(x, 1) - d(y, 1)|$ , as stated.

(ii)  $\Rightarrow$  (iii) Take  $x \geq y \geq z$  and let  $N(x) = d(x, 1)$ . Then

$$\begin{aligned} d(x, y) + d(y, z) &= |N(x) - N(y)| + |N(y) - N(z)| \\ &= N(y) - N(x) + N(z) - N(y) = N(z) - N(x) \\ &= |N(x) - N(z)| = d(x, z). \end{aligned}$$

The strict monotonicity of  $d(1, x)$  follows from the symmetry of  $d$  and the fact that  $d(x, y) = |d(1, x) - d(1, y)| = 0$  if and only if  $x = y$ .

(iii)  $\Rightarrow$  (i) Define  $N(x) = d(1, x)$ . First of all, since  $d(1, x)$  is strictly monotone,  $d(1, 1) = 0$  and  $d(1, 0) = 1$ , it follows that  $N$  is a strict negation. Moreover, if  $1 \geq y \geq z$ , it follows that

$$d(1, y) + d(y, z) = d(1, z)$$

so  $d(y, z) = d(1, z) - d(1, y)$ , and since  $y > z$ ,  $d(y, z) = |d(1, y) - d(1, z)| = |N(y) - N(z)| \square$

On the other hand, we can also state the following result.

**Theorem 4** Let  $d_R : [0, 1]^2 \rightarrow [0, 1]$  be a restricted dissimilarity function. Then  $d_R$  is quasi-convex in one variable; that is, for all  $x, y_1, y_2, \lambda \in [0, 1]$

$$d_R(x, \lambda y_1 + (1 - \lambda)y_2) \leq \max(d_R(x, y_1), d_R(x, y_2)) \quad (1)$$

*Proof.* Take  $x, \lambda \in [0, 1]$ . We know that, for all  $y_1, y_2 \in [0, 1]$

$$\min(y_1, y_2) \leq \lambda y_1 + (1 - \lambda)y_2 \leq \max(y_1, y_2) \quad (2)$$

There are three possibilities.

i)  $x \leq \min(y_1, y_2)$ . Then

$$\begin{aligned} d_R(x, \lambda y_1 + (1 - \lambda)y_2) &\leq d_R(x, \max(y_1, y_2)) \\ &\leq \max(d_R(x, y_1), d_R(x, y_2)) \end{aligned}$$

ii)  $\min(y_1, y_2) \leq x \leq \max(y_1, y_2)$  In this situation two things can happen:

a)  $\min(y_1, y_2) \leq x \leq \lambda y_1 + (1 - \lambda)y_2 \leq \max(y_1, y_2)$ , so  $d_R(x, \lambda y_1 + (1 - \lambda)y_2) \leq d_R(x, \max(y_1, y_2)) \leq \max(d_R(x, y_1), d_R(x, y_2))$ , or

b)  $\lambda y_1 + (1 - \lambda)y_2 \leq x \leq \max(y_1, y_2)$ , so  $d_R(x, \lambda y_1 + (1 - \lambda)y_2) \leq d_R(x, \min(y_1, y_2)) \leq \max(d_R(x, y_1), d_R(x, y_2))$  iii)  $\max(y_1, y_2) \leq x$ , can be treated as item i)  $\square$

**Corollary 1** Let  $d_R : [0, 1]^2 \rightarrow [0, 1]$  be a restricted dissimilarity function. If  $\lambda_0, \lambda_1 \in [0, 1]$ , then

$$\begin{aligned} d_R(\lambda_0 x_1 + (1 - \lambda_0)x_2, \lambda_1 y_1 + (1 - \lambda_1)y_2) &\leq \\ &\max(d_R(x_1, y_1), d_R(x_1, y_2), d_R(x_2, y_1), d_R(x_2, y_2)) \end{aligned}$$

for all  $x_1, x_2, y_1, y_2 \in [0, 1]$ .

*Proof.* From the symmetry of  $d_R$  we have that

$$d_R(\lambda_0 x_1 + (1 - \lambda_0)x_2, \lambda_1 y_1 + (1 - \lambda_1)y_2)$$

is smaller than or equal to

$$\begin{aligned} &\max(d_R(\lambda_0 x_1 + (1 - \lambda_0)x_2, y_1), \\ &d_R(\lambda_0 x_1 + (1 - \lambda_0)x_2, y_2)) \end{aligned}$$

which is also less than or equal to

$$\begin{aligned} &\max(\max(d_R(x_1, y_1), d_R(x_2, y_1)), \\ &\max(d_R(x_1, y_2), d_R(x_2, y_2))). \end{aligned}$$

The following example shows that a restricted dissimilarity function needs not be quasi-convex in both of its arguments.

**Example 2** Take  $d_R(x, y) = (|x^2 - y^2|)^{\frac{1}{2}}$ ,  $\lambda = \frac{1}{2}$ ,  $x_1^2 = \frac{1}{4}$ ,  $x_2^2 = 0$ ,  $y_1^2 = \frac{3}{4}$  and  $y_2^2 = \frac{1}{2}$ . Then calculation shows that  $d_R(\lambda x_1 + (1 - \lambda)y_1, \lambda x_2 + (1 - \lambda)y_2) = \frac{1}{2}|\frac{1}{2} + \frac{\sqrt{3}}{2}|^{\frac{1}{2}} > \frac{1}{2}$  whereas  $d_R(x_1, x_2) = \frac{1}{2} = d_R(y_1, y_2)$ .

Restricted dissimilarity functions which are concave can be related to metrics as follows.

**Theorem 5** Let  $d_R[0, 1]^2 \rightarrow [0, 1]$  be a restricted dissimilarity function which is concave in each co-ordinate. Then  $d_R$  is a metric on  $[0, 1]$ .

*Proof.* We have to check the triangle inequality of  $d_R$  only, as the other properties of metrics are trivially fulfilled by  $d_R$ . The only non-trivial case to be checked is when  $0 \leq x < z < y \leq 1$ . Then the concavity in one coordinate ensures for each  $\lambda \in [0, 1]$  that

$$\begin{aligned} d_R(x, \lambda x + (1 - \lambda)y) &\geq \lambda d_R(x, x) + (1 - \lambda)d_R(x, y) \\ &= (1 - \lambda)d_R(x, y). \end{aligned}$$

Similarly,

$$d_R(\lambda x + (1 - \lambda)y, y) \geq \lambda d_R(x, y).$$

Thus, taking  $\lambda = \frac{y-z}{y-x}$

$$\begin{aligned} d_R(x, y) &= \frac{y-z}{y-x} d_R(x, y) + \frac{z-x}{y-x} d_R(x, y) \\ &\leq d_R(x, z) + d_R(y, z), \end{aligned}$$

proving that  $d_R$  is a metric

**Remark**

- (i)  $d_R$  satisfying Theorem 5 is necessarily strict.
- (ii)  $d_R(x, y) = (x - y)^2$  is a strict dissimilarity restricted function which is not a metric (note that it is not concave in each coordinate).

Analogously to the case of penalty functions we can introduce the concept of faithful restricted dissimilarity function as follows.

**Definition 7** [3] A faithful restricted dissimilarity function is a restricted dissimilarity function  $d_R$  such that there exist a strictly increasing and continuous function  $h : [0, 1] \rightarrow [0, 1]$  and a convex

function  $K : [-1, 1]^2 \rightarrow [0, 1]$  with a unique minimum at  $K(0) = 0$  such that

$$d_R(x, y) = K(h(x) - h(y))$$

for every  $x, y \in [0, 1]$ .

We can provide a characterization of restricted dissimilarity functions in terms of automorphisms and bijections, as follows.

**Proposition 1** [3] For any mapping  $d_R : [0, 1]^2 \rightarrow [0, 1]$  the following statements are equivalent:

1.  $d_R$  is a faithful restricted dissimilarity function.
2. There exist a convex automorphism  $\varphi : [0, 1] \rightarrow [0, 1]$  and a continuous bijection  $h : [0, 1] \rightarrow [0, 1]$  such that

$$d_R(x, y) = \varphi(|h(x) - h(y)|)$$

for every  $x, y \in [0, 1]$ .

**Example 3** Take  $\alpha \in ]0, \infty]$ . If we consider  $\varphi(x) = x^2$  and  $h(x) = x^\alpha$  we see that the mapping

$$d_R(x, y) = |x^\alpha - y^\alpha|^2$$

is a faithful restricted dissimilarity function.

#### 4. Construction of penalty functions by means of convex functions

Faithful restricted dissimilarity functions can be used to build penalty functions as the following result shows.

**Theorem 6** [3] Let  $d_R : [0, 1]^2 \rightarrow [0, 1]$  be a faithful restricted dissimilarity function and  $(w_1, \dots, w_n) \in [0, 1]^n$  such that  $\sum_{i=1}^n w_i = 1$ . Then the function

$$P(x_1, \dots, x_n, y) = \sum_{i=1}^n w_i d_R(x_i, y)$$

is a penalty function.

Note that if we define the aggregation function:

$$M(x_1, \dots, x_n) = \sum_{i=1}^n w_i x_i$$

which is a weighted mean, the previous penalty function can be written as:

$$P(x_1, \dots, x_n, y) = M(d_R(x_1, y), \dots, d_R(x_n, y)) .$$

This expression raises the question of whether other types of aggregation functions and/or other kinds of dissimilarity-like functions can be used to obtain penalty functions. In this sense, a first result is the following, which shows how convex aggregation functions and convex dissimilarity-like functions can be used for such construction.

**Proposition 2** Let  $d_i : [0, 1]^2 \rightarrow [0, 1]$ ,  $i = 1, 2, \dots, n$  be a family of functions which are convex in their second variable and such that  $d_i(x, y) = 0$  if  $x = y$ . Let  $M : [0, 1]^n \rightarrow [0, 1]$  be a convex aggregation function. Then the function  $P : [0, 1]^{n+1} \rightarrow [0, 1]$  defined by:

$$P(x_1, \dots, x_n, y) = M(d_1(x_1, y), \dots, d_n(x_n, y))$$

is a penalty function.

*Proof.*

1. If  $x_i = y$  for every  $i = 1, \dots, n$ , then from the properties demanded to  $d_i$  and the fact that  $M(0, \dots, 0) = 0$  it follows that  $P(x, \dots, x, x) = 0$ .
2. Quasi-convexity of  $P$  follows from a straight calculation, since the composition of a convex non-decreasing function with convex functions is also convex.

In particular, this result implies that we need to deal with convex aggregation functions. In this sense, note that for instance the maximum and the weighted mean are convex. However, the study of convexity for general aggregation functions is not easy, so it would be desirable to find an alternative approach.

#### 5. Construction of penalty functions by means of maxitive and quasi-convex functions

Notice that, in fact, we do not need to make use of full convexity. Quasi-convex dissimilarity-like functions and maxitive aggregation functions are sufficient for building penalty functions.

The concept of maxitivity was introduced by Dubois and Prade [8] in order to aggregate possibility measures. It reads as follows.

**Definition 8** A function  $M : [0, 1]^n \rightarrow [0, 1]$  is maxitive if:

$$\begin{aligned} & M(\max(x_1, y_1), \dots, \max(x_n, y_n)) \\ &= \max(M(x_1, \dots, x_n), M(y_1, \dots, y_n)) \end{aligned}$$

for every  $(x_1, \dots, x_n), (y_1, \dots, y_n) \in [0, 1]^n$ .

$M$  is submaxitive if

$$\begin{aligned} & M(\max(x_1, y_1), \dots, \max(x_n, y_n)) \\ &\leq \max(M(x_1, \dots, x_n), M(y_1, \dots, y_n)) \end{aligned}$$

for every  $(x_1, \dots, x_n), (y_1, \dots, y_n) \in [0, 1]^n$ .

**Example 4** The function

$$M(x_1, \dots, x_n) = \max_{i=1}^n x_i^{w_i} \text{ with } w_i > 0$$

is an example of maxitive aggregation function.

The following result completely characterizes maxitive and submaxitive aggregation functions.

**Proposition 3** *Let  $M : [0, 1]^n \rightarrow [0, 1]$  be an aggregation function. Then*

1.  *$M$  is maxitive if and only if there exist non-decreasing functions  $f_i : [0, 1] \rightarrow [0, 1]$ ,  $i = 1 \dots, n$  with  $f_i(0) = 0$  for any  $i = 1 \dots, n$  and  $\max_{i=1}^n f_i(1) = 1$  such that:*

$$M(x_1, \dots, x_n) = \max_{i=1}^n f_i(x_i)$$

*for every  $(x_1, \dots, x_n) \in [0, 1]^n$ .*

2.  *$M$  is sub-maxitive if and only if it is maxitive.*

*Proof.*

1. See [8] (or [9] p. 53).
2. It is straight since from the monotonicity of aggregation functions it always holds that:

$$\begin{aligned} & M(\max(x_1, y_1), \dots, \max(x_n, y_n)) \\ & \geq \max(M(x_1, \dots, x_n), M(y_1, \dots, y_n)) \end{aligned}$$

for every  $(x_1, \dots, x_n), (y_1, \dots, y_n) \in [0, 1]^n$ .

Now we can state our main result.

**Theorem 7** *Let  $d_i : [0, 1]^2 \rightarrow [0, 1]$ ,  $i = 1, \dots, n$  be functions which are quasi-convex in their second variable and such that  $d_i(x, y) = 0$  whenever  $x = y$ . Let  $M : [0, 1]^n \rightarrow [0, 1]$  be a submaxitive (and hence maxitive) aggregation function. Then, the function  $P : [0, 1]^{n+1} \rightarrow [0, 1]$ , defined for any  $x_1, \dots, x_n, y \in [0, 1]$  as:*

$$P(x_1, \dots, x_n, y) = M(d_1(x_1, y), \dots, d_n(x_n, y))$$

*is a penalty function.*

*Proof.* Only quasi-convexity in  $y$  requires some consideration. For the sake of simplicity we consider the case  $n = 2$ , the general case being analogous. So take  $x_1, x_2, y_1, y_2, \lambda \in [0, 1]$ . We have that

$$M(d_1(x_1, \lambda y_1 + (1 - \lambda)y_2), d_2(x_2, \lambda y_1 + (1 - \lambda)y_2))$$

is less than or equal to

$$M(\max(d_1(x_1, y_1), d_1(x_1, y_2)), \max(d_2(x_2, y_1), d_2(x_2, y_2)))$$

due to the quasi-convexity of  $d_i$ . But, since  $M$  is submaxitive, this expression is smaller than or equal to

$$\max(M(d_1(x_1, y_1), d_2(x_2, y_1)), M(d_1(x_1, y_2), \dots, d_2(x_2, y_2)))$$

as we wanted to see.

Regarding dissimilarity-like functions which are quasi-convex in their second variable, observe that a particular instance is that of restricted dissimilarity functions. More generally, we can state the following.

**Proposition 4** *Let  $d : [0, 1]^2 \rightarrow [0, 1]$  be a commutative function such that  $d(x, y) \leq d(x, z)$  and  $d(y, z) \leq d(x, z)$  whenever  $x \leq y \leq z$ . Then  $d$  is quasi-convex in its second variable.*

*Proof.* Given  $x, y_1, y_2, \lambda \in [0, 1]$ , we have to prove that  $d(x, \lambda y_1 + (1 - \lambda)y_2) \leq \max(d(x, y_1), d(x, y_2))$ . There are two possible cases:

1.  $x \leq \lambda y_1 + (1 - \lambda)y_2$ . Since  $\lambda y_1 + (1 - \lambda)y_2 \leq \max(y_1, y_2)$ , we get

$$d(x, \lambda y_1 + (1 - \lambda)y_2) \leq d(x, \max(y_1, y_2))$$

But it always holds that  $d(x, \max(y_1, y_2)) \leq \max(d(x, y_1), d(x, y_2))$ . So from both inequalities we arrive at:

$$d(x, \lambda y_1 + (1 - \lambda)y_2) \leq \max(d(x, y_1), d(x, y_2))$$

2.  $x \geq \lambda y_1 + (1 - \lambda)y_2$ . It is analogous, just taking into account that  $\min(y_1, y_2) \leq \lambda y_1 + (1 - \lambda)y_2$ .

## 6. Conclusions and future research

In this work we have presented a method that allows to build penalty functions by means of appropriate aggregation and restricted dissimilarity functions. Analyzing the properties that are necessary for such construction, we have also been able to provide a method for constructing penalty functions in a similar way but with functions that are only requested to be quasi-convex in their second variable.

At the same time, we have studied some properties of restricted dissimilarity functions. We have related them to metrics. In future works we intend to extend this analysis to the dissimilarity like functions.

## Acknowledgements

The authors were partially supported by the Spanish projects TIN2012-32482, TIN2010-15055 and TEC2012-39095-C03-02, and by grants APVV-0073-10 and GACR P 402/11/0378 .

## References

- [1] G. Beliakov, H. Bustince, D. Paternain, Image Reduction Using Means on Discrete Product Lattices. *IEEE Transactions on Fuzzy Systems* **21** (2012) 1070-1083.
- [2] G. Beliakov, A. Pradera and T. Calvo, *Aggregation Functions: A Guide for Practitioners* Springer, Berlin, Heidelberg, 2007.
- [3] H. Bustince, A. Jurio, A. Pradera, R. Mesiar, G. Beliakov. Generalization of the weighted voting method using penalty functions constructed via faithful restricted dissimilarity functions. *European Journal of Operational Research* **225** (2013) 472-478.

- [4] H. Bustince, E. Barrenechea and M. Pagola, Relationship between restricted dissimilarity functions, restricted equivalence functions and normal EN-functions: Image thresholding invariant, *Pattern Recognition Letters*, **29**, (4), (2008) 525-536.
- [5] T. Calvo, A. Kolesarova, M. Komornikova and R. Mesiar, Aggregation operators: properties, classes and construction methods. In T. Calvo, G. Mayor and R. Mesiar (Eds.): *Aggregation Operators New Trends and Applications* (Physica-Verlag, Heidelberg, 2002); 3 - 104.
- [6] T. Calvo and G. Beliakov, Aggregation functions based on penalties, *Fuzzy Sets and Systems*, 161 (2010), 1420-1436.
- [7] T. Calvo, R. Mesiar and R. Yager, A quantitative weights and aggregation, *IEEE Transactions on Fuzzy Systems*, **12**, (2004), 62 - 69.
- [8] D. Dubois and H. Prade, Aggregation of Possibility Measures, In J. Kacprzyk, M. Fedrizzi (eds): *Multiperson Decision Making Models Using Fuzzy Sets and Possibility Theory* (Kluwer Academic Publishers, 1990), 55-63.
- [9] M. Grabisch, J.L. Marichal, R. Mesiar and E. Pap. Aggregation Functions, Cambridge University press, (2009).
- [10] A.N. Kolmogoroff, Sur la notion de la moyenne. *Accad. Naz. Lincei Mem. Cl. Sci. Fis. Mat. Natur. Sez.* **12** (1930), 388 - 391.
- [11] R. Mesiar, Fuzzy set approach to the utility, preference relations, and aggregation operators, *European Journal of Operational Research* **176** (2007) 414 - 422.
- [12] M. Nagumo, Über eine Klasse der Mittelwerte. *Japanese Journal of Mathematics* **6** (1930), 71 - 79.