Dissimilarity, Quasimetric, Metric

Ehtibar N. Dzhafarov
Purdue University and Swedish Collegium for Advanced Study

Abstract

Dissimilarity is a function that assigns to every pair of stimuli a nonnegative number vanishing if and only if two stimuli are identical, and that satisfies the following two conditions called the intrinsic uniform continuity and the chain property, respectively: it is uniformly continuous with respect to the uniformity it induces, and, given a set of stimulus chains (finite sequences of stimuli), the dissimilarity between their initial and terminal elements converges to zero if the chains' length (the sum of the dissimilarities between their successive elements) converges to zero. The four properties axiomatizing this notion are shown to be mutually independent. Any conventional, symmetric metric is a dissimilarity function. A quasimetric (satisfying all metric axioms except for symmetry) is a dissimilarity function if and only if it is symmetric in the small. It is proposed to reserve the term metric (not necessarily symmetric) for such quasimetrics. A real-valued binary function satisfies the chain property if and only if whenever its value is sufficiently small it majorates some quasimetric and converges to zero whenever this quasimetric does. The function is a dissimilarity function if, in addition, this quasimetric is a metric with respect to which the function is uniformly continuous.

Keywords: asymmetric metric, categorization, discrimination, dissimilarity, Fechnerian Scaling, metric, quasimetric, stimulus space, subjective metric.

This note aims at filling in certain conceptual and terminological gaps in the theory of dissimilarity as presented in Dzhafarov and Colonius (2007) and Dzhafarov (2008a, b). The axioms defining dissimilarity are proved to be mutually independent. A convenient criterion is given for the compliance of a function with the chain property ($\mathcal{D}4$ below). Psychophysical applications of the notion of dissimilarity were previously confined to discrimination judgments, "same-different" and "greater-less." Here, an example is given of an application of the notion to the categorization paradigm. As a terminological improvement, it is suggested that the term "oriented" (i.e., asymmetric) metric should be reserved for quasimetrics (functions satisfying all metric axioms except for symmetry) which are symmetric in the small: then any metric is a dissimilarity function. A familiarity with the dissimilarity cumulation theory and its main psychophysical application, Fechnerian Scaling (at least as they are presented in Dzhafarov & Colonius, 2007), is desirable for understanding the context of this note.

Notation conventions. Let $\mathfrak{S}$ be a set *stimuli* (*points*), denoted by boldface lowercase letter $\mathbf{x}, \mathbf{y}, \ldots$. A *chain*, denoted by boldface capitals, $\mathbf{X}, \mathbf{Y}, \ldots$, is a finite sequence of points. The set $\bigcup_{k=0}^{\infty} \mathfrak{S}^k$ of all chains with elements in $\mathfrak{S}$ is denoted by $\mathcal{S}$. It contains the empty chain and one-element chains (identified with their elements, so that $\mathbf{x} \in \mathfrak{S}$ is also the chain consisting of $\mathbf{x}$). Concatenations of two or more chains are presented by concatenations of their symbols, $\mathbf{XY}$, $\mathbf{xYz}$, etc. Binary real-valued functions $\mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ are presented as $D\mathbf{xy}$, $M\mathbf{xy}$, $\gamma\mathbf{xy}$, ...

Given a chain $\mathbf{X} = \mathbf{x}_1, \ldots \mathbf{x}_n$ and a binary (real-valued) function F, the notation $F\mathbf{X}$ stands for

$$\sum_{i=1}^{n-1} F\mathbf{x}_i\mathbf{x}_{i+1},$$

with the obvious convention that the quantity is zero if n is 1 (one-element chain) or 0 (empty chain). The notation $\Delta_F \mathbf{X}$ stands for

$$\max_{i=1,\ldots,n-1} F\mathbf{x}_i\mathbf{x}_{i+1}.$$

All limit statements for sequences, such as $F\mathbf{a}_n\mathbf{a} \to 0$, $F\mathbf{a}_n\mathbf{b}_n \to 0$, or $F\mathbf{X}_n \to 0$ are implicitly predicated on $n \to \infty$.

If stimuli $\mathbf{x}, \mathbf{y}$, etc. are characterized by numerical values of a certain property (e.g., auditory tones are characterized by their intensity), these numerical values may be denoted by x, y, etc. (possibly with indices or ornaments), and by abuse of notation we can say $\mathbf{x} = x$, $\mathbf{y} = y$, etc. Analogously, if the stimuli are characterized by numerical vectors, we can write $\mathbf{x} = (x_1, \ldots x_k)$.

1. Dissimilarity: Definition and Applications

1.1 Definition. Function $D : \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ is a *dissimilarity function* if it has the following properties:

$\mathcal{D}1$(*positivity*) $D\mathbf{ab} > 0$ for any distinct $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$;

$\mathcal{D}2$ (*zero property*) $D\mathbf{aa} = 0$ for any $\mathbf{a} \in \mathfrak{S}$;

$\mathcal{D}3$ (*intrinsic uniform continuity*) for any $\varepsilon > 0$ one can find a $\delta > 0$ such that, for any $\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}' \in \mathfrak{S}$,

$$\text{if } D\mathbf{aa}' < \delta \text{ and } D\mathbf{bb}' < \delta, \text{ then } |D\mathbf{a}'\mathbf{b}' - D\mathbf{ab}| < \varepsilon;$$

$\mathcal{D}4$ (*chain property*) for any $\varepsilon > 0$ one can find a $\delta > 0$ such that for any chain $\mathbf{aXb}$,

$$\text{if } D\mathbf{aXb} < \delta, \text{ then } D\mathbf{ab} < \varepsilon.$$

1.2 Remark. The intrinsic uniform continuity and chain properties can also be presented in terms of sequences, as, respectively,

$$\text{if } (D\mathbf{a}_n\mathbf{a}'_n \to 0) \text{ and } (D\mathbf{b}_n\mathbf{b}'_n \to 0) \text{ then } D\mathbf{a}'_n\mathbf{b}'_n - D\mathbf{a}_n\mathbf{b}_n \to 0,$$

and

$$\text{if } D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0 \text{ then } D\mathbf{a}_n\mathbf{b}_n \to 0,$$

for all sequences of points $\{\mathbf{a}_n\}$, $\{\mathbf{a}'_n,\}$ $\{\mathbf{b}_n\}$, $\{\mathbf{b}'_n\}$ in $\mathfrak{S}$ and sequences of chains $\{\mathbf{X}_n\}$ in $\mathcal{S}$. (For the equivalence of these statements to $\mathcal{D}3$ and $\mathcal{D}4$ see Dzhafarov & Colonius, 2007, footnote 8.)

1.3 Remark. A simple but useful observation: in the formulation of $\mathcal{D}4$ "for any chain $\mathbf{aXb}$" can be replaced with "for any chain $\mathbf{aXb}$ with $D\mathbf{aXb} < m$," where m is any positive number. Indeed, the δ to be found for a given ε can always be redefined as $\min\{\delta, m\}$. In this class of chains we also have $\Delta\mathbf{aXb} < m$, where Δ stands for Δ_D (see Notation Conventions above). Stated in terms of sequences, if

$$\text{if } D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0 \text{ then } \Delta\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0,$$

so one can only consider the sequences $\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$ with sufficiently small $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$ and $\Delta\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n$.

The convergence $D\mathbf{x}_n\mathbf{y}_n \to 0$ is easily shown to be an equivalence relation on the set of all infinite sequences in $\mathfrak{S}$ (Dzhafarov & Colonius, 2007, Theorem 1).

1.4 Theorem. *Each of the properties $\mathcal{D}1 - \mathcal{D}4$ is independent of the remaining three.*

Proof. (1) Let $\mathfrak{S}$ be a two-element set $\{\mathbf{a}, \mathbf{b}\}$. If

$$D\mathbf{ab} = D\mathbf{ba} = D\mathbf{aa} = D\mathbf{bb} = 0,$$

then $\mathcal{D}1$ is violated while $\mathcal{D}2 - \mathcal{D}4$ hold trivially.

(2) With the same $\mathfrak{S}$, if

$$D\mathbf{ab} = D\mathbf{ba} = D\mathbf{aa} = D\mathbf{bb} = 1,$$

then $\mathcal{D}2$ is violated while $\mathcal{D}1, \mathcal{D}3, \mathcal{D}4$ hold trivially.

(3) Let $\mathfrak{S}$ be represented by $\mathbb{R}$, and let

$$D\mathbf{xy} = \begin{cases} |x-y| & \text{if } |x-y| < 1 \\ |x-y| + \mathrm{e}^{|x-y|} - \mathrm{e} & \text{if } |x-y| \geq 1 \end{cases}.$$

Then D obviously satisfies $\mathcal{D}1 - \mathcal{D}2$. To see that $\mathcal{D}4$ holds too, we use Remark 1.3 and only consider chains $\mathbf{aXb}$ with $\Delta\mathbf{aXb} \leq D\mathbf{aXb} < 1$. Since $D\mathbf{aXb} \geq |a-b|$, we have then $D\mathbf{ab} = |a-b| < \varepsilon$ whenever $D\mathbf{aXb} < \delta = \varepsilon$. But D does not satisfy $\mathcal{D}3$, which can be shown by choosing, e.g., $a = a' = 0$ and $b' = b + \delta$ ($b > 1$, $\delta > 0$): then

$$D\mathbf{a}'\mathbf{b}' - D\mathbf{ab} = \delta + \mathrm{e}^{b}\left(\mathrm{e}^{\delta} - 1\right)$$

considered as a function of b can be arbitrarily large for any choice of $\delta = D\mathbf{bb}' > D\mathbf{aa}' = 0$.

(4) Let $\mathfrak{S}$ be represented by a finite interval of reals and $D\mathbf{xy} = (x-y)^2$. Then D satisfies $\mathcal{D}1 - \mathcal{D}3$ (obviously) but not $\mathcal{D}4$: subdivide a nondegenerate interval $[a, b]$ within $\mathfrak{S}$ into n equal parts and observe that the length of the corresponding chain is

$$D\mathbf{aXb} = n\left(\frac{a-b}{n}\right)^2 \to 0$$

as $n \to \infty$, while $D\mathbf{ab}$ is fixed at $(a-b)^2 > 0$. □

The following three examples show how the notion of dissimilarity appears in the contexts of perceptual discrimination and categorization.

1.5 Example. Let the stimulus set $\mathfrak{S}$ be represented by $[1, U]$ (e.g., the set of intensities of sound between an absolute threshold taken for 1 and an upper threshold U), and let $\gamma\mathbf{xy}$ be the probability with which $\mathbf{y}$ is judged to be greater than $\mathbf{x}$ in some respect (e.g., loudness). Let

$$\gamma\mathbf{xy} = \Phi\left(\frac{y-x}{x}\right),$$

where Φ is a continuously differentiable probability distribution function on $\mathbb{R}$ with $\Phi(0) = \frac{1}{2}$ and $\Phi'(0) > 0$. Then the quantity

$$D\mathbf{xy} = \left|\Phi\left(\frac{y-x}{x}\right) - \frac{1}{2}\right|$$

is a dissimilarity function.[1] Indeed, its compliance with $\mathcal{D}2$ is clear, and $\mathcal{D}1$ follows from the fact that Φ is nondecreasing on $\mathbb{R}$ and increasing in an open neighborhood of zero. To demonstrate the intrinsic uniform continuity, $\mathcal{D}3$, observe that the convergence $D\mathbf{x}_n\mathbf{x}'_n \to 0$ on $[1, U] \times [1, U]$ is equivalent to the conventional convergence $|x_n - x'_n| \to 0$, and that on this compact area $D\mathbf{xy}$ is uniformly continuous in the conventional sense. Finally, to demonstrate the chain property, $\mathcal{D}4$, we use Remark 1.3 to conclude that $D\mathbf{a}_n\mathbf{X}_n\mathbf{b}_n \to 0$ implies (assuming, without loss of generality, $a_n \leq b_n$)

$$\int_{a_n}^{b_n} \left|\Phi\left(\frac{\mathrm{d}x}{x}\right) - \frac{1}{2}\right| = \int_{a_n}^{b_n} \left|\Phi'(0)\frac{\mathrm{d}x}{x}\right| = \Phi'(0)\log\frac{b_n}{a_n} \to 0,$$

and the latter is equivalent to

$$D\mathbf{a}_n\mathbf{b}_n = \frac{b_n - a_n}{a_n} \to 0.$$

1.6 Example. Consider a set of probability vectors

$$\mathfrak{S}_0 = \left\{(p_1, \ldots, p_k) \in [0,1]^k : \sum p_i = 1\right\}$$

[1]This is a special case of the "mirror-reflection" procedure described in Dzhafarov and Colonius (1999).

The value of p_i can be interpreted as the probability with which a stimulus is classified into the ith category among an exhaustive list of k mutually exclusive categories. Two stimuli corresponding to one and the same vector $(p_1, \ldots, p_k)$ are treated as indistinguishable, so one can speak of the vector as a stimulus. Define a measure of "difference" of a point $\mathbf{q} = (q_1, \ldots, q_k)$ from a point $\mathbf{p} = (p_1, \ldots, p_k)$ as[2]

$$D\mathbf{pq} = \sqrt{\sum_{i=1}^{k} (q_i - p_i) \log \frac{q_i}{p_i}}.$$

Then, for any $c \in \left]0, \frac{1}{k}\right[$, D is a (symmetric) dissimilarity function on

$$\mathfrak{S}_c = \left\{ (p_1, \ldots, p_k) \in [c, 1 - (k-1)\,c]^k : \sum p_i = 1 \right\}.$$

The properties $\mathcal{D}1 - \mathcal{D}2$ are verified trivially. The intrinsic uniform continuity, $\mathcal{D}3$, follows from the fact that D is uniformly continuous in the conventional sense on the compact domain $\mathfrak{S}_c \times \mathfrak{S}_c$, and that the conventional convergence (in Euclidean norm) $|\mathbf{p}_n - \mathbf{p}'_n| \to 0$ is equivalent to $D\mathbf{p}_n\mathbf{p}'_n \to 0$. (Note that this will no longer be true if we replace $p_i \geq c$ with $p_i \geq 0$ or $p_i > 0$ for some $i \in \{1, \ldots, k\}$.) To see that D satisfies the chain property, $\mathcal{D}4$, we use the inequality

$$D\mathbf{pq} \geq V\mathbf{pq} = \sum_{i=1}^{k} |p_i - q_i|,$$

an immediate consequence of Pinsker's inequality (Csiszàr, 1967).[3] It follows that for any chain $\mathbf{X}$ with elements in $\mathfrak{S}_c$,

$$D\mathbf{pXq} \geq V\mathbf{pq},$$

whence the convergence $D\mathbf{p}_n\mathbf{X}_n\mathbf{q}_n \to 0$ implies $V\mathbf{p}_n\mathbf{q}_n \to 0$, which in turn is equivalent to $D\mathbf{p}_n\mathbf{q}_n \to 0$.

1.7 Example. Let $\mathfrak{S}$ be any finite set of points, and $D\mathbf{xy}$ any function subject to $\mathcal{D}1 - \mathcal{D}2$. Then $\mathcal{D}3$ and $\mathcal{D}4$ are satisfied trivially, and D is a dissimilarity function.

2. Conditions Relating Dissimilarities, Quasimetrics, and Metrics

Of the four properties defining dissimilarity, the chain property, $\mathcal{D}4$, is the only one not entirely intuitive. It is not always obvious how one should go about testing the compliance of a function with this property. The criterion (necessary and sufficient condition) given in Theorem 2.5 below often proves helpful in this issue. Conveniently and somewhat surprisingly, it turns out to be a criterion for the conjunction of the conditions $\mathcal{D}1, \mathcal{D}2, \mathcal{D}4$ rather than just $\mathcal{D}4$, and it may even help in dealing with $\mathcal{D}3$.

We need first to establish terminological clarity in treating metric-like functions which may lack symmetry. By simply dropping the symmetry axiom from the list of metric axioms one creates the notion of a quasimetric.

2.1 Definition. Function $M : \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ is a *quasimetric function* if it has the following properties:

$\mathcal{QM}1$(*positivity*) $M\mathbf{ab} > 0$ for any distinct $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$;

$\mathcal{QM}2$ (*zero property*) $M\mathbf{aa} = 0$ for any $\mathbf{a} \in \mathfrak{S}$;

$\mathcal{QM}3$ (*triangle inequality*) $M\mathbf{ab} + M\mathbf{bc} \geq M\mathbf{ac}$ for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{S}$.

---

[2] $D^2\mathbf{pq}$ is the symmetric Kullback-Leibler divergence, the original divergence measure proposed in Kullback and Leibler (1951). The present example does not touch on many important aspects of the relationship between divergence measures in the information geometry and the dissimilarity cumulation theory: this is a topic for a separate treatment.

[3] The familiar form of Pinsker's inequality is $K^2\mathbf{pq} = \sum_{i=1}^{k} q_i \log \frac{q_i}{p_i} \geq \frac{1}{2} V^2\mathbf{pq}$, from which the present form obtains by $D^2\mathbf{pq} = K^2\mathbf{pq} + K^2\mathbf{qp}$.

Any symmetric metric is obviously a quasimetric. If M is a quasimetric then $M\mathbf{xy}+M\mathbf{yx}$ (or $\max\{M\mathbf{xy}, M\mathbf{yx}\}$) is a symmetric metric.

The notion of a quasimetric, however, turns out to be too unconstrained to serve as a good formalization for the intuition of an asymmetric ("oriented") metric. In particular, a quasimetric does not generally induce a uniformity,[4] which means that generally it is not a dissimilarity function (see Dzhafarov & Colonius, 2007).

2.2 Example. Let $\mathfrak{S}$ be represented by the interval $]0,1[$, and let

$$Q\mathbf{xy} = H_0\,(x-y) + (y-x)\,,$$

where $H_0\,(a)$ is a version of the Heaviside function (equal 0 for $a \le 0$ and equal 1 for $a > 0$). The function clearly satisfies $\mathcal{QM}1 - \mathcal{QM}2$, and $\mathcal{QM}3$ follows from

$$H_0\,(u) + H_0\,(v) \ge H_0\,(u+v)\,,$$

for any real u, v. But Q is not a dissimilarity function as it is not intrinsically uniformly continuous: take, e.g., any $a = b = b'$ and let $a'_n \to a+$ in the conventional sense. Then $Q\mathbf{a}\mathbf{a}'_n = a'_n - a \to 0$ and $Q\mathbf{b}\mathbf{b}' = 0$, but $Q\mathbf{a}'_n\mathbf{b}' = 1 + b' - a'_n \to 1$ while $Q\mathbf{ab} = 0$. [Compared to the uniformity induced by a dissimilarity function (Dzhafarov & Colonius, 2007, Section 2.2), one can check that the sets $\mathfrak{A}_\varepsilon = \{(\mathbf{x},\mathbf{y}) : Q\mathbf{xy} < \varepsilon\}$ taken for all positive ε do not form a uniformity basis on $\mathfrak{S}$ because, for any $\varepsilon < 1$ and any $\delta > 0$, $\mathfrak{A}_\varepsilon^{-1} = \{(\mathbf{y},\mathbf{x}) : Q\mathbf{xy} < \varepsilon\}$ does not include as a subset $\mathfrak{A}_\delta$.]

We see that simply dropping the symmetry axiom is not completely innocuous. At the same time the symmetry requirement is too stringent both in and without the present context (e.g., in differential geometry symmetry is usually unnecessary in treating intrinsic metrics). A satisfactory definition of an asymmetric metric can be achieved by replacing the symmetry requirement with "symmetry in the small."

2.3 Definition. Function $M : \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ is a *metric (or distance function)* if it is a quasimetric with the following property:

$\mathcal{M}$(symmetry in the small) for any $\varepsilon > 0$ one can find a $\delta > 0$ such that $M\mathbf{ab} < \delta$ implies $M\mathbf{ba} < \varepsilon$, for any $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$.

What is important for our purposes is that any metric in the sense of this definition is a dissimilarity function.

2.4 Theorem. *A quasimetric is a dissimilarity function if and only if it is a metric.*

Proof. A dissimilarity function satisfies the symmetry in the small condition (Dzhafarov & Colonius, 2007, Theorem 1). This proves the "only if" part. For the "if" part, the compliance of a metric M with $\mathcal{D}1 - \mathcal{D}2$ is obvious. By the triangle inequality, for any $\mathbf{a}, \mathbf{b}, \mathbf{a}', \mathbf{b}' \in \mathfrak{S}$,

$$\begin{cases} M\mathbf{a}\mathbf{a}' + M\mathbf{b}'\mathbf{b} & \ge M\mathbf{ab} - M\mathbf{a}'\mathbf{b}' \\ M\mathbf{a}'\mathbf{a} + M\mathbf{b}\mathbf{b}' & \ge M\mathbf{a}'\mathbf{b}' - M\mathbf{ab} \end{cases}.$$

From Definition 2.3, for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\max\{M\mathbf{a}'\mathbf{a}, M\mathbf{b}'\mathbf{b}\} < \frac{\varepsilon}{2}$$

whenever

$$\max\{M\mathbf{a}\mathbf{a}', M\mathbf{b}\mathbf{b}'\} < \min\left\{\frac{\varepsilon}{2}, \delta\right\} \le \frac{\varepsilon}{2}.$$

This proves $\mathcal{D}3$, as it follows that $|M\mathbf{ab} - M\mathbf{a}'\mathbf{b}'|$ can be made less than ε. The chain property, $\mathcal{D}4$, follows from $M\mathbf{aXb} \ge M\mathbf{ab}$, for all $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$ and $\mathbf{X} \in \mathcal{S}$, which holds by the triangle inequality. □

[4] For the notion of uniformity see, e.g., Kelly (1955, Chapter 6). A brief reminder of the basic properties of a uniformity and the relation of this notion to that of a dissimilarity function can be found in Dzhafarov and Colonius (2007, Sections 2.2 and 2.6).

The intrinsic metrics of differential geometry always satisfy Definition 2.3, and so does the metric

$$M\mathbf{ab} = \inf_{\mathbf{X}\in\mathcal{S}} D\mathbf{aXb}$$

induced by a dissimilarity function D (see Dzhafarov & Colonius, 2007, Theorems 7 and 9). It seems that most of the basic results pertaining to metrics can be obtained with metrics in the sense of Definition 2.3 (see Dzhafarov, 2008a, b, for the demonstration of this statement on the notion of path length).[5]

We are now prepared to formulate a criterion for the chain property of a dissimilarity function, followed by a sufficient condition for intrinsic uniform continuity.

2.5 Theorem. *A function $D : \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ satisfies $\mathcal{D}1$, $\mathcal{D}2$, and $\mathcal{D}4$ if and only if, for some $m \in]0, \infty]$ and any $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$,*

$$\text{if } D\mathbf{ab} < m \text{ then } D\mathbf{ab} \geq M\mathbf{ab},$$

where $M : \mathfrak{S} \times \mathfrak{S} \to \mathbb{R}$ is a quasimetric with the following property: for any $\varepsilon > 0$ one can find a $\delta_\varepsilon > 0$ such that, for any $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$,

$$\text{if } M\mathbf{ab} < \delta_\varepsilon, \text{ then } D\mathbf{ab} < \varepsilon.$$

2.6 Remark. A slightly less rigorous but more compact formulation of the theorem is this: a real-valued binary function D on $\mathfrak{S}$ satisfies $\mathcal{D}1$, $\mathcal{D}2$, and $\mathcal{D}4$ if and only if for all $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$ with sufficiently small $D\mathbf{ab}$ the latter majorates a quasimetric $M\mathbf{ab}$ such that $M\mathbf{a}_n\mathbf{b}_n \to 0$ implies $D\mathbf{a}_n\mathbf{b}_n \to 0$.

Proof. If such a quasimetric M exists, D satisfies $\mathcal{D}1$ because if $\mathbf{a} \neq \mathbf{b}$, then either $D\mathbf{ab} \geq m > 0$ or $D\mathbf{ab} \geq M\mathbf{ab} > 0$. $\mathcal{D}2$ follows from the observations that (1) for any $\varepsilon > 0$ we have $0 = M\mathbf{aa} < \delta_\varepsilon$, implying $D\mathbf{aa} < \varepsilon$; and (2) choosing $\varepsilon < m$ we have $D\mathbf{aa} \geq M\mathbf{aa} = 0$. To show the compliance of D with $\mathcal{D}4$, note that for any $\varepsilon > 0$, if $D\mathbf{aXb} < \min\{m, \delta_\varepsilon\}$ then

$$D\mathbf{aXb} \geq M\mathbf{aXb} \geq M\mathbf{ab} < \delta_\varepsilon,$$

implying $D\mathbf{ab} < \varepsilon$.

Conversely, if D satisfies $\mathcal{D}1$, $\mathcal{D}2$, and $\mathcal{D}4$, then function

$$M\mathbf{ab} = \inf_{\mathbf{X}\in\mathcal{S}} D\mathbf{aXb}$$

is a quasimetric. Its compliance with $\mathcal{QM}3$ follows from

$$M\mathbf{ab} + M\mathbf{bc} = \inf_{\mathbf{X},\mathbf{Y}\in\mathcal{S}} D\mathbf{aXbYc} = \inf_{\substack{\mathbf{Z}\in\mathcal{S}\\ \mathbf{b} \text{ is in } \mathbf{Z}}} D\mathbf{aZc} \geq \inf_{\mathbf{Z}\in\mathcal{S}} D\mathbf{aZc} = M\mathbf{ac}.$$

To see that M satisfies $\mathcal{QM}2$ and $\mathcal{QM}1$, observe that $D\mathbf{aXb} \geq 0$, for all $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$ and $\mathbf{X} \in \mathcal{S}$. Since $\mathcal{S}$ includes the empty chain, we get

$$M\mathbf{aa} = \inf_{\mathbf{X}\in\mathcal{S}} D\mathbf{aXa} = 0.$$

For $\mathbf{a} \neq \mathbf{b}$,

$$M\mathbf{ab} = \inf_{\mathbf{X}\in\mathcal{S}} D\mathbf{aXb} > 0,$$

because if the infimum could be zero then, for every $\delta > 0$, we would be able to find an $\mathbf{X}$ with $D\mathbf{aXb} < \delta$, and this would contradict $\mathcal{D}4$ since $D\mathbf{ab} > 0$. To complete the proof it remains to observe that $D\mathbf{ab} \geq M\mathbf{ab}$ for all $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$. □

[5] In the previous publications on dissimilarity and Fechnerian Scaling the term "oriented metric" was used to designate what we presently propose to call simply "metric" (adding "symmetric" if dealing with a conventional metric). Nor was in the previous publications the notion of an asymmetric metric differentiated from that of a quasimetric with sufficient clarity (although Theorem 2.4 is mentioned in Dzhafarov & Colonius, 2007, Section 2.1).

In practice the quasimetric M of the theorem often can be chosen to be a metric in the sense of Definition 2.3. This makes it possible to speak of the function D as being or not being uniformly continuous with respect to (the uniformity induced by) the metric M: D possesses this property if $M\mathbf{a}_n\mathbf{a}'_n \to 0$ and $M\mathbf{b}_n\mathbf{b}'_n \to 0$ imply $D\mathbf{a}'_n\mathbf{b}'_n - D\mathbf{a}_n\mathbf{b}_n \to 0$.

2.7 Corollary. *If M in Theorem 2.5 is a metric, then D and M induce the same uniformity, that is, the convergence $D\mathbf{x}_n\mathbf{y}_n \to 0$ is equivalent to the convergence $M\mathbf{x}_n\mathbf{y}_n \to 0$, for any sequences $\{\mathbf{x}_n\}, \{\mathbf{y}_n\}$ in $\mathfrak{S}$. Consequently, if D is uniformly continuous with respect to (the uniformity induced by) M, then D is a dissimilarity function.*

Proof. The implication "if $M\mathbf{x}_n\mathbf{y}_n \to 0$ then $D\mathbf{x}_n\mathbf{y}_n \to 0$" holds by the definition of M, while the reverse implication follows from D majorating M whenever the former is sufficiently small. As a result, if D is uniformly continuous with respect to M then it satisfies $\mathcal{D}3$ (and by the previous theorem it satisfies $\mathcal{D}1, \mathcal{D}2, \mathcal{D}4$). □

Thus, the properties $\mathcal{D}1$, $\mathcal{D}2$, and $\mathcal{D}4$ in Example 1.5 could be established by noting that for a sufficiently small $m > 0$, $\left|\Phi(a) - \frac{1}{2}\right|$ majorates $k|a|$ on the interval $\left[\Phi^{-1}\left(m - \frac{1}{2}\right), \Phi^{-1}\left(m + \frac{1}{2}\right)\right]$, where k is any positive number less than the minimum $\Phi'(a)$ on this interval. It would follow then that if $D\mathbf{xy} = \left|\Phi\left(\frac{y-x}{x}\right) - \frac{1}{2}\right| < m$ (where $x, y \in [1, U]$), then

$$D\mathbf{xy} > k\frac{|y-x|}{x} \geq \frac{k}{U}|y-x| = M\mathbf{xy},$$

the latter being a (symmetric) metric. Since $M\mathbf{x}_n\mathbf{y}_n \to 0$ clearly implies $D\mathbf{x}_n\mathbf{y}_n \to 0$, the properties in question follow by Theorem 2.5. The property $\mathcal{D}3$ obtains by Corollary 2.7: M and D are uniformly equivalent (i.e., one of them converges to zero if and only if so does the other), and it is clear that D is uniformly continuous on $[1, U]$ in the conventional sense (which here means, with respect to M).

In Example 1.6, Pinsker's inequality is all one needs to mention to establish $\mathcal{D}1$, $\mathcal{D}2$, and $\mathcal{D}4$ by Theorem 2.5. $\mathcal{D}3$ follows by Corollary 2.7 using the same reasoning as in the previous example.

Acknowledgement

This research has been supported by NSF grant SES 0620446 and AFOSR grants FA9550-06-1-0288 and FA9550-09-1-0252 to Purdue University. The author is most grateful to Hans Colonius for critically reading a draft of this note and suggesting improvements.

References

Csiszàr, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, 2, 299–318.

Dzhafarov, E. N. (2008a). Dissimilarity cumulation theory in arc-connected spaces. *Journal of Mathematical Psychology*, 52, 73–92. [See also Dzhafarov, E. N. (2009). Corrigendum to: "Dissimilarity cumulation theory in arc-connected spaces"' [J. Math. Psychol. 52 (2008) 73–92]. *Journal of Mathematical Psychology*, 53, 300.]

Dzhafarov, E. N. (2008b). Dissimilarity cumulation theory in smoothly-connected spaces. *Journal of Mathematical Psychology*, 52, 93–115.

Dzhafarov, E. N., & Colonius, H. (1999). Fechnerian metrics in unidimensional and multidimensional stimulus spaces. *Psychonomic Bulletin and Review*, 6, 239–268.

Dzhafarov, E. N., & Colonius, H. (2007). Dissimilarity cumulation theory and subjective metrics. *Journal of Mathematical Psychology*, 51, 290–304.

Kelly, J.L. (1955). *General Topology.* Toronto: Van Nostrand.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency, *Annals of Mathematical Statistics*, 22, 79–86.