ULTRAMETRIC FECHNERIAN SCALING OF DISCRETE OBJECT SETS

Hans Colonius, Ehtibar N. Dzhafarov

Abstract: *Universal Fechnerian Scaling (UFS) is a principled approach to computing "subjective" distances among objects (stimuli) from their pairwise discrimination probabilities. It is based on the concept of 'dissimilarity function' leading to a locally symmetrical quasimetric in the context of Dissimilarity Cumulation (DC) theory developed by Dzhafarov and Colonius. Here we show that, for finite sets of objects, the replacement of dissimilarity cumulation with a dissimilarity maximization procedure results in "subjective" distances satisfying the ultrametric inequality.*

Keywords: *Fechnerian scaling; dissimilarity function; quasimetric; ultrametric.*

ACM Classification Keywords*: G.2.3 Discrete Mathematics – Applications;*

MSC*: 54E05, 54E15, 54E35, 05C12*

Introduction

In experimental psychology, there are many different procedures for gaining insight into a person's subjective representation of a given set of objects, like colors or tones, pictures of faces, semantic terms, etc. Eliciting numerical judgments about the perceived similarity of a pair of stimuli (objects), or estimating the probability with which a stimulus $\mathbf{x}$ is recognized as stimulus $\mathbf{y}$ (confusion probabilities), are among the most common methods with a long tradition in this field [Shepard, 1957]. A discrimination paradigm focal for this paper involves a set of stimuli $\mathfrak{S} = \{\mathbf{s_1}, \mathbf{s_2}, \ldots, \mathbf{s_N}\}$, $N > 1$, presented two at a time to a perceiver whose task is to respond to each ordered pair $(\mathbf{x}, \mathbf{y})$ by "$\mathbf{x}$ and $\mathbf{y}$ are the same" or by "$\mathbf{x}$ and $\mathbf{y}$ are different". Each ordered pair $(\mathbf{x}, \mathbf{y})$ is then assigned (an estimate of) the discrimination probability function[1]

$$\psi\mathbf{xy} = \Pr\left[\mathbf{x} \text{ and } \mathbf{y} \text{ are judged to be different}\right] \tag{1}$$

Often the investigator seeks an embedding of subjects' judgments in a Euclidian (or, Minkowskian) space via multidimensional scaling (MDS), or in an ultrametric space (rooted tree structure) via cluster analysis (CA), such that distances among the points representing the stimuli correspond as closely as possible to the observed discrimination probabilities. Both procedures are based on the *probability-distance hypothesis* [Dzhafarov, 2002], that is, the assumption that for some distance function $H\mathbf{xy}$ and some increasing transformation f

$$\psi\mathbf{xy} = f(H\mathbf{xy}). \tag{2}$$

The problem for MDS and CA is that experimental data show systematic violations of *symmetry* and *constant self-dissimilarity* (i.e., generally, $\psi\mathbf{xy} = \psi\mathbf{yx}$ and $\psi\mathbf{xx} = \psi\mathbf{yy}$, respectively, for distinct $\mathbf{x}$ and $\mathbf{y}$).

Dzhafarov and Colonius [Dzhafarov, Colonius, 2007] developed a principled approach to solve this problem, *Universal Fechnerian Scaling* (UFS), which is applicable to all possible (finite or infinite) stimulus spaces endowed with "same-different" discrimination probabilities. UFS is based on the theory of *Dissimilarity Cumulation* (DC) [Dzhafarov, Colonius, 2007; Dzhafarov, 2008a; Dzhafarov, 2008b], which provides a general definition of a *dissimilarity function* and shows how to impose topological and metric properties on stimulus sets.

[1] Notation convention: real-valued functions of one or more arguments that are elements of a stimulus set are indicated by strings without parentheses, e.g., $\psi\mathbf{xy}$ instead of $\psi(\mathbf{x}, \mathbf{y})$

The next section gives an outline of only those aspects of UFS and DC theory that are relevant for finite object sets. For the general case, we refer to the papers above and [Dzhafarov, 2010a]. Subsequently, a new (finite) variant based on the ultrametric property will be introduced.

Concepts and Results from DC Theory and UFS for Finite Spaces

Notation conventions. Let $\mathfrak{S}$ be a finite set of objects (stimuli). A *chain*, denoted by boldface capitals, $\mathbf{X}, \mathbf{Y}, \ldots$, is a finite sequence of objects. 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 (conveniently 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. 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).

Dissimilarity function and quasimetric. For a finite set $\mathfrak{S}$, a real-valued function $D : \mathfrak{S} \times \mathfrak{S} \rightarrow \Re$ 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 $\mathbf{a} \in \mathfrak{S}$.

Note that a dissimilarity function need not be symmetric and need not satisfy the triangle inequality. A dissimilarity function M that does satisfy the triangle inequality is called a *quasimetric*:

$$M\mathbf{ab} + M\mathbf{bc} \geq M\mathbf{ac} \tag{3}$$

for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{S}$.

Definition 1. Given a dissimilarity D on a finite set $\mathfrak{S}$, the **quasimetric G induced by D** is defined as

$$G\mathbf{ab} = \min_{\mathbf{X} \in \mathcal{S}} D\mathbf{aXb}, \tag{4}$$

for all $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$.

That G is a quasimetric is easy to prove (see, e.g., [Dzhafarov, Colonius, 2006; Dzhafarov, Colonius, 2007]).

Psychometric increments and Fechnerian distance. If $\mathfrak{S}$ is endowed with ψ, as defined in (1), then (following a certain "canonical" transformation of the stimuli)[2], the dissimilarity function can be defined as either of the two kinds of *psychometric increments*

$$\Psi^{(1)}\mathbf{ab} = \psi\mathbf{ab} - \psi\mathbf{aa} \;\; \text{and} \;\; \Psi^{(2)}\mathbf{ab} = \psi\mathbf{ba} - \psi\mathbf{aa}. \tag{5}$$

Due to the canonical form of ψ, these quantities are always positive for $\mathbf{b} \neq \mathbf{a}$. Denoting by D either $\Psi^{(1)}$ or $\Psi^{(2)}$, one uses (4) to compute, respectively, the quasimetrics $G_1\mathbf{ab}$ or $G_2\mathbf{ab}$ (called *Fechnerian distances*).

The quantity

$$G^*\mathbf{ab} = G_1\mathbf{ab} + G_1\mathbf{ba} = G_2\mathbf{ab} + G_2\mathbf{ba} \tag{6}$$

then is a metric on $\mathfrak{S}$, called the *overall Fechnerian distance*. The equality of the two computations is easy to establish [Dzhafarov, Colonius, 2006].

[2] For details, including the concept of *regular minimality* see [Dzhafarov, Colonius, 2006]

Ultrametric Fechnerian Scaling of Discrete Object Sets

Any empirical data set being finite, DC theory for finite sets can be viewed as an alternative data-analytic tool to nonmetric (MDS): rather than seeking a nonlinear transformation of a given set of dissimilarities into a (usually Euclidean) metric, DC replaces the dissimilarity value for each ordered pair of points with the shortest length ("cumulated dissimilarity") of a finite chain of points connecting the first element of the pair with the second one resulting in a quasimetric induced by the dissimilarity function.

However, for many empirical dissimilarities on sets of high-dimensional stimuli, or with an underlying collection of hierarchical features, a representation by a rooted tree structure may sometimes be more appropriate. Thus, it seems reasonable to look for a data-analytic alternative to CA based on the principles of Fechnerian Scaling. The basic idea consists in replacing the "dissimilarity cumulation" procedure by "dissimilarity maximization". Given a chain $\mathbf{X} = \mathbf{x}_1, \ldots \mathbf{x}_n$ and a binary (real-valued) function F, the notation $\Delta_F\mathbf{X}$ stands for

$$\max_{i=1,\ldots,n-1} F\mathbf{x_i x_{i+1}},$$

again with the obvious convention that the quantity is zero if n is 1 or 0. A dissimilarity function M on a finite set $\mathfrak{S}$ is called a *quasi-ultrametric* if it satisfies the ultrametric inequality,

$$\max\{M\mathbf{ab}, M\mathbf{bc}\} \geq M\mathbf{ac} \tag{7}$$

for all $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathfrak{S}$.

Definition 2. Given a dissimilarity D on a finite set $\mathfrak{S}$, the *quasi-ultrametric G^∞ induced by D* is defined as

$$G^\infty\mathbf{ab} = \min_{\mathbf{X}\in\mathcal{S}} \Delta_D\mathbf{aXb}, \tag{8}$$

for all $\mathbf{a}, \mathbf{b} \in \mathfrak{S}$.

That G^∞ is a quasi-ultrametric is easy to prove. A reasonable symmetrization procedure, yielding a metric that can be called the *overall Fechnerian ultrametric*, is

$$G^{\infty^*}\mathbf{ab} = \max\{G^\infty\mathbf{ab}, G^\infty\mathbf{ba}\} \tag{9}$$

then yields a (symmetric) ultrametric on $\mathfrak{S}$, called the *overall Fechnerian ultrametric*.

Conclusion and Further Developments

On a finite set, any dissimilarity function induces a quasimetric by the "dissimilarity cumulation" procedure of DC theory [Dzhafarov, Colonius, 2006; Dzhafarov, Colonius, 2007]. Here we have suggested a complementary approach inducing a quasi-ultrametric by a "dissimilarity maximization" procedure. A systematic comparison between the two procedures for different types of dissimilarity data sets is left for future investigation, but the following can be noted right away.

First, unlike the DC theory, its ultrametric counterparts does not generalize to arbitrary stimulus sets. Thus, for arc-connected spaces the Fechnerian ultrametric is identically equal to zero.

Second, unlike in the UFS, the overall Fechnerian ultrametric is not the same for the two kinds of psychometric increments, as defined in (5): the equality of the two sums in (6) does not have an analog with (9) in which G^∞ is computed from $\Psi^{(1)}$ and from $\Psi^{(2)}$.

There is, however, one important similarity. [Dzhafarov, 2010b] has shown that the procedure of computing quasimetric distances from dissimilarities can also be described in terms of a series of recursive corrections of the dissimilarity values for violations of the triangle inequality. It can be shown that a corresponding series of recursive corrections on the dissimilarity values for violations of the ultrametric inequality would yield the induced quasi-ultrametric distances.

One can consider procedures intermediate between cumulation and maximization of dissimilarities by defining, for any dissimilarity function D, the length of a chain $\mathbf{X} = \mathbf{x}_1, \ldots \mathbf{x}_n$ by

$$D\mathbf{X} = ((D\mathbf{x_1 x_2})^k + \ldots + (D\mathbf{x}_{n-1}\mathbf{x}_n)^k)^{1/k}. \tag{10}$$

For $k \to \infty$ this would result in the ultrametric approach outlined above. For finite k, the procedure is generalizable to arbitrary dissimilarity spaces. This follows from the fact the use of (10) is equivalent to the use of the original DC in which one, first, redefines D into D^k (which yields another dissimilarity function), and then redefines the locally symmetrical quasimetric G induced by D^k into $G^{1/k}$ (which yields another locally symmetrical quasimetric).

Acknowledgements

H.C. acknowledges support by DFG (Grant SFB-TRR31 Active Listening). E.D. acknowledges support by AFOSR grant FA9550-09-1-0252.

Bibliography

[Dzhafarov, 2002] E.N.Dzhafarov. Multidimensional Fechnerian scaling: Probability-distance hypothesis. J Math Psychol, vol. 46, 352-374

[Dzhafarov, 2008a] E.N.Dzhafarov. Dissimilarity cumulation theory in arc-connected spaces. J Math Psychology, vol. 52, 73-92.

[Dzhafarov, 2008b] E.N.Dzhafarov. Dissimilarity cumulation theory in smoothly-connected spaces. J Math Psychology, vol. 52, 93-115.

[Dzhafarov, 2010a] E.N.Dzhafarov. Dissimilarity, quasidistance, distance. J Math Psychology, vol 54, 334-337.

[Dzhafarov, 2010b] E.N.Dzhafarov. Dissimilarity Cumulation as a procedure correcting for violations of triangle inequality. J Math Psychology, vol. 54, 284-287.

[Dzhafarov, Colonius, 2006] E.N.Dzhafarov and H.Colonius. Reconstructing distances among objects from their discriminability. Psychometrika, vol. 71, 365-386.

[Dzhafarov, Colonius, 2007] E.N.Dzhafarov and H.Colonius. Dissimilarity cumulation theory and subjective metrics, J of Math Psychology, vol. 51(5), 290-304.

[Shepard, 1957] R.N.Shepard. Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space, Psychometrika, vol. 22, 325-345, 1957.

Authors' Information

Hans Colonius *- Carl von Ossietzky Universitaet Oldenburg, 26111 Oldenburg, Germany;*
e-mail: hans.colonius@uni-oldenburg.de
Major Fields of Scientific Research: Mathematical modeling of elementary cognitive processes, Psychophysics

Ehtibar N. Dzhafarov *- Purdue University, Department of Psychological Sciences,*
703 Third Street, West Lafayette, IN 47907-2081, USA; e-mail: ehtibar@purdue.edu
Major Fields of Scientific Research: Mathematical psychology, Psychophysics