A Topological Definition of Phase and Amplitude Variability of Functional Data

Simone Vantini

Key words: functional data analysis, phase variability, amplitude variability, registration, alignment, synchronization, warping

1 Introduction

In functional data analysis (e.g., [2]), the problem of data registration is encountered when the variability between subjects - represented as functions $y_i(x)$ with $i = 1, \dots, n$ - is assumed to be related not only to the dependent variable y but also to the independent one x; in a very wide sense, registering functional data means identifying this second source of variability and removing it by means of subject-dependent suitable transformations of the independent variable. Functional data registration is often a necessary step to achieve a successful functional data analysis since it allows a correct matching across subjects. More technically, a registration of a functional data set is considered any procedure that aims at making the n observed functions y_i as similar as possible by means of n suitable transformations of the abscissas h_i. These transformations are commonly named warping functions and the variability of the functional data set imputable to them is usually named *phase variability*; finally, the residual variability observed among the aligned functions is named *amplitude variability*. In the recent literature, the spread use of continuous alignment procedures in applications coexists with a very little number of theoretical works trying to formalize the problem of curve registration (e.g., [1]). In this work, we propose a possible mathematical framework where this problem can be coherently set: we show indeed that the introduction, in a functional data analysis, of a metric and of a group of warping functions, respect to which the metric

Simone Vantini
MOX - Department of Mathematics, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milan, Italy, e-mail: simone.vantini@polimi.it

is invariant, enables a sound and not ambiguous definition of phase and amplitude variability. Indeed in this framework, we prove that the analysis of a continuously-registered functional data set can be re-interpreted as the analysis of a set of suitable equivalence classes associated to original functions and induced by the group of the warping functions.

2 Functional Data Registration Revisited

To obtain this coherent formalization of the problem of registration some basic properties, of the set F which the functional data belong to, and of the set W of warping functions, are demanded:

a) $F = \{f : \Omega \subseteq \mathbf{R}^p \longrightarrow \Psi \subseteq \mathbf{R}^q\}$ is a metric space according to a metric $d : F \times F \longrightarrow \mathbf{R}_0^+$,
b) W is a subgroup – with respect to ordinary composition $\circ$ – of the group of the continuous automorphisms: $\Omega \subseteq \mathbf{R}^p \longrightarrow \Omega \subseteq \mathbf{R}^p$;
c) $\forall f \in F$ and $\forall h \in W$ we have that $f \circ h \in F$;
d) Given any couple of elements $f_1, f_2 \in F$ and an element $h \in W$, the distance between f_1 and f_2 is invariant under the composition of f_1 and f_2 with h, i.e.:

$$d(f_1, f_2) = d(f_1 \circ h, f_2 \circ h) \; ; \tag{1}$$

we will refer to this property as W*-invariance* of d.

Thanks to properties (a)-(d), it is possible to define a semi-metric $d_W : F \times F \longrightarrow \mathbf{R}_0^+$ that is jointly determined by the metric d and the group W ([3]):

Theorem 1 (Definition of the semi-metric d_W).
Under properties (a)-(d), $d_W(f_1, f_2) := \min_{h_1,h_2 \in W} d(f_1 \circ h_1, f_2 \circ h_2)$ *, when defined, is a semi-metric.*

Sufficient conditions for the existence of the minimum are reported in [3]. Like any other semi-metric, d_W induced a partition of the space F in to a quotient set that we will indicate as $\mathscr{F}$. The W-invariance of the original metric d provides a bijective correspondence between the equivalence classes of the quotient set $\mathscr{F}$ and the orbits of the action of the group W on the set F. Thus, we can define a metric $d_{\mathscr{F}} : \mathscr{F} \times \mathscr{F} \longrightarrow \mathbf{R}_0^+$ on $\mathscr{F}$ that is consistent with the original metric d on F ([3]); let $[f]$ indicate the equivalence class of $\mathscr{F}$ which f belongs to:

Theorem 2 (Definition of the metric $d_{\mathscr{F}}$).
Under properties (a)-(d), $d_{\mathscr{F}}([f_1],[f_2]) := d_W(f_1, f_2)$ *, is a metric.*

We are now going to formalize in this mathematical framework the problem of the registration of a couple of functions f_1 and f_2.

Definition 1. Given f_1 and $f_2 \in F$ and a minimizing couple h_1 and $h_2 \in W$ (i.e. h_1 and h_2 such that $d(f_1 \circ h_1, f_2 \circ h_2) = d_{\mathscr{F}}([f_1],[f_2])$), $\tilde{f_1} = f_1 \circ h_1$ and $\tilde{f_2} = f_2 \circ h_2$ are said *mutually-registered representatives* of $[f_1]$ and $[f_2]$.

Note that, since given a couple of elements f_1 and $f_2 \in F$ there is not a unique minimizing couple h_1 and h_2, there is not a unique couple $\tilde{f}_1$ and $\tilde{f}_2$ of mutually-registered representatives of $[f_1]$ and $[f_2]$. It is worth mentioning two certain special couples of mutually-registered representatives of $[f_1]$ and $[f_2]$: the one corresponding to $h_1 = \mathbf{1}$ and the one corresponding to $h_2 = \mathbf{1}$. In the former case $\tilde{f}_1 = f_1$ while in the latter case $\tilde{f}_2 = f_2$.

Definition 2. Given a couple f_1 and f_2, and a couple of mutually-registered representatives $\tilde{f}_1$ and $\tilde{f}_2$ such that $\tilde{f}_1 = f_1$ and $h_1 = \mathbf{1}$, $\tilde{f}_2$ is said a f_1*-registered representative* of $[f_2]$ (or in less formal but more familiar terms $\tilde{f}_2$ is said a *registered* version of f_2 with respect to f_1). We will refer to it as $\tilde{f}_{2\to 1}$ and to the corresponding warping function as $h_{2\to 1}$. The definition of $\tilde{f}_{1\to 2}$ and $h_{1\to 2}$ is analogous.

Moreover, since an f_1-registered representative of $[f_2]$ is an element $\in [f_2]$ minimizing the distance with f_1, we have that:

$$\tilde{f}_{2\to 1} = \arg\min_{f\in[f_2]} d(f, f_1) \quad \text{and} \quad \tilde{f}_{1\to 2} = \arg\min_{f\in[f_1]} d(f, f_2) \ ,$$

with $d(\tilde{f}_{1\to 2}, f_2) = d(f_1, \tilde{f}_{2\to 1}) = d_{\mathscr{F}}([f_1],[f_2])$. According to this framework, registering a function $f_1 \in F$ with respect to a function $f_2 \in F$ - according to a metric d and a class of warping functions W - simply means replacing f_1 with $\tilde{f}_{1\to 2}$.

The introduction of a quotient set $\mathscr{F}$ over F (dependent on the choices for d and W) is the key to a clear and not ambiguous definition of *Phase Variability* and *Amplitude Variability*. We are quite sure to meet the heuristic sense of many authors, by defining the phase variability as the one that can occur between functions $\in F$ belonging to the same equivalence class, i.e. the variability within equivalence classes (in this case, $d_{\mathscr{F}}([f_1],[f_2]) = 0$). Coherently, the amplitude variability is the variability between functions not belonging to the same equivalence class and not imputable to phase variability, i.e. the variability between equivalence classes (in this case we have $d_{\mathscr{F}}([f_1],[f_2]) = d(f_1, f_2)$).

In the same framework, it is straightforward to define the registration of a set $\{f_i\}_{i=1,2,\dots,n}$ with respect to a target function f_0. Indeed registering the set $\{f_i\}_{i=1,2,\dots,n}$ with respect to f_0 means replacing the set $\{f_i\}_{i=1,2,\dots,n}$ with the set $\{\tilde{f}_{i\to 0}\}_{i=1,2,\dots,n}$ (or simply $\{\tilde{f}_i\}_{i=1,2,\dots,n}$) whose distances to f_0 are minimal over the relevant equivalence classes:

$$\{f_i\}_{i=1,2,\dots,n} \longmapsto \{\tilde{f}_i = \arg\min_{f\in[f_i]} d(f_0, f)\}_{i=1,2,\dots,n} \ .$$

In other words, registering the set $\{f_i\}_{i=1,2,\dots,n}$ with respect to f_0 consists in finding in $[f_1]$, $[f_2]$, ..., $[f_n]$, n functions that are the closest to f_0 respectively.

Given the fact that $0 \le d_{\mathscr{F}}([f_1],[f_2]) \le d(f_1, f_2)$, we can define an amplitude-to-total variability ratio bounded between 0 and 1, useful in practical situations, measuring to what extent phase and amplitude variability contribute to total variability:

$$\alpha^2 = \frac{\sum_{i=1}^n d^2_{\mathscr{F}}([f_i],[f_0])}{\sum_{i=1}^n d^2(f_i, f_0)} \ ;$$

and then we can then simply characterize the two extreme situations as follows:

- presence of phase variability only, when $\alpha^2 = 0$;
- presence of amplitude variability only, when $\alpha^2 = 1$.

The two extreme situations can be equivalently characterized as follows:

- presence of phase variability only, when for $i = 1,2,\ldots,n : \tilde{f}_i \equiv f_0$;
- presence of amplitude variability only, when for $i = 1,2,\ldots,n : \tilde{f}_i \equiv f_i$.

Note that both the registration of a set of functions in absence of a target function and the registration of a set of functions when d is a semi-metric can also be inserted in this framework; because of the limited space, these issues will be not addressed here. To this regard, please refer to [3].

3 Conclusions and Future Perspectives

On the whole, by introducing the semi-metric d_W, we managed to formalize the problem of registration by showing that performing an analysis of a functional data set using a semi-metric d_W is either equivalent to perform an analysis of suitable equivalence classes using the metric $d_{\mathscr{F}}$ (i.e., our theoretical abstraction) and equivalent to perform an analysis of the registered functions using the original metric d (i.e., the usual approach used in applications).

In the present paper we also propose an amplitude-to-total variability ratio α^2 that is here used as a purely descriptive tool. It is of paramount interest to investigate in the future the possibility of using it as an inferential tool for testing the absence/presence of phase variability. Identifying the distribution a suitable test statistic derived from α^2 under the null hypothesis of absence of phase variability will require the introduction of a probabilistic model for both amplitude and phase variability and, probably, also an “anova-inspired” decomposition of total variability in amplitude and phase variability.

Acknowledgements We want to thank Laura Maria Sangalli and Piercesare Secchi for having stimulated this work: the idea of a coherence between the metric and the class of warping functions – which this work starts from – emerged while working with them within the AneuRisk Project.

References

1. Davidian, M., Lin, X., Wang, J. (2004): *Introduction [Emerging issues in longitudinal and functional data analysis]*, Statistica Sinica - special issue, Vol. 14, n. 3, pp. 613–614.
2. Ramsay, J.O., Silverman, B.W. (2005): *Functional data analysis*. Second edition. Springer Series in Statistics. Springer, New York.
3. Vantini, S. (2012): *On the Definition of Phase and Amplitude Variability in Functional Data Analysis*, Test, in press.