Rejoinder
Functional Modeling and Classification of Longitudinal Data

HANS-GEORG MÜLLER

University of California, Davis

Ivar Heuch and Rima Izem emphasize the need to take up new challenges facing statistics by increasingly complex applied problems. FDA provides an as of yet incomplete toolbox for a class of such challenges. The observation of Ivar Heuch that the practical role of functional principal components in applied multivariate analysis is similar to that in functional data analysis implies that the accumulated knowledge about interpreting principal components and the techniques that were developed in multivariate analysis should be brought to full bearing in FDA. This comment is also echoed by Jim Ramsay in the specific context of rotating principal components to enhance their interpretation, a time-honored tool of applied multivariate analysis. Classical multivariate technology clearly plays a very important role in FDA and I wholeheartedly agree that the experience and applied methodology that has been accumulated in multivariate analysis over several decades should be fully exploited in FDA.

It is important to complement functional principal components, given their role as a fundamental tool in FDA, by nonlinear alternatives, as Rima Izem points out. Besides the nonlinear modes of variation, other extensions may look at modifying the principle of minimizing orthogonal least squares, which is one way to characterize principal components, by allowing for more flexible metrics, such as incorporating L^1 distances in the functional situation. Rima Izem's observation about the important role of graphical output is very well taken; in a way, plotting eigenfunctions is more natural than trying to visualize, say, a 10-dimensional principal component vector. This is a consequence of the powerful concepts of time order, continuity and smoothness that come into play in FDA, but do not matter for multivariate analysis.

Jim Ramsay directs our attention to the fact that by now there already exist various approaches to achieve the necessary regularization task in FDA. Functional principal components as a tool for dimension reduction have the advantage that they seem to be overall well interpretable (an excellent discussion is provided by Jones & Rice, 1992, who demonstrate how to explore functional data in various eigen-directions), easily modifiable and implementable with a variety of smoothing methods, ranging from kernel smoothers and local linear fitting to various splines and wavelets. The smoothing and multivariate analysis concepts are thus disentangled, which I view as an advantage. Last not least, functional principal components seems to be the only approach to date for which notable foundations have been laid and some ground walls erected for a yet to be completed edifice that might eventually constitute an asymptotic theory and methodology of inference for FDA. Future promise may also be

found in yet another approach, the purely "nonparametric" methodology for dimension reduction in FDA, which is still in its infancy.

As Rima Izem notes, there is a clear need for the development of appropriate functional methods for objects such as shapes and images. I could not agree more in directing attention to the analysis of such objects that are currently not very standard topics within FDA. Analyzing image data with functional methods seems promising. Issues such as image warping become quite difficult when subjected to a rigorous analysis (Glasbey & Mardia, 2001) and the systematic use of functional concepts may lead to progress. Other sometimes non-standard data structures for which FDA may prove very useful arise in genomics and the analysis of microarrays.

It would be no bad thing if statisticians were to acquire more boldness in attacking seemingly non-standard scientific data and their underlying problems with the versatile tools and methodologies they have at their disposal. Such boldness and the broad perspective that goes with it is commonly encountered among computer scientists in the field of machine learning. Support vector machines and other learning methods are being applied to classify text documents, complex biological information, trees and other highly non-standard objects (Shawe-Taylor & Cristianini, 2004). Developing similarly bold extensions and applications of functional data analysis and the related complex statistical methodology to encompass such non-standard data structures is a valuable goal for future research.

Additional References

Glasbey, C.A. & Mardia, K.V. (2001). A penalized likelihood approach to image warping. *J. Royal Statist. Soc., Series B*, **63**, 465-514.

Jones, M.C. & Rice, J. (1992). Displaying the important features of large collections of similar curves. *Amer. Statist.*, **46**, 140-145.

Shawe-Taylor, J. & Cristianini, N. (2004). Kernel methods for pattern analysis. Cambridge University Press.