RESEARCH BRIEF

How can depth of data help solve a neuroimaging problem?

Myriam Vimond, Pavlo Mozharovskyi, Pierre Lafaye de Micheaux

December 2020

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

How can depth of data help solve a neuroimaging problem ?

Pierre Lafaye de Micheaux, Pavlo Mozharovskyi, Myriam Vimond

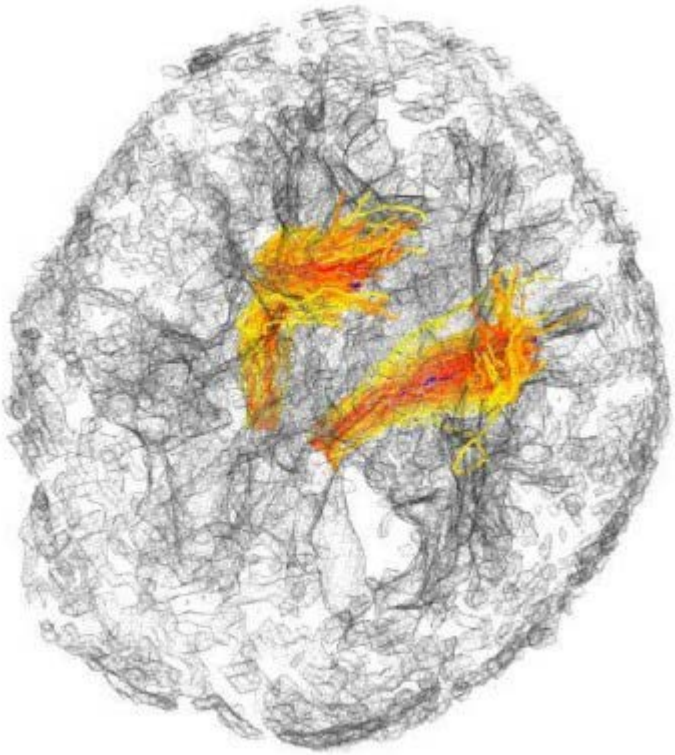
Does genetics play a role in the spatial organization of neuronal fibers?
This approach was originally motivated by the need to analyze a very large number of white matter fiber bundles obtained by diffusion tensor imaging (an MRI-based neuroimaging technique) in a population of elderly twins. How can the white fibers of a brain be analyzed to detect possible abnormalities that could lead to degenerative diseases such as Alzheimer's? Are these diseases only hereditary or can external factors cause the brain to lose white fibers?

The article by Pierre Lafaye de Micheaux, Pavlo Mozharovskyi & Myriam Vimond, «Depth for Curve Data and Applications», is at the crossroads of the three co-authors' practices (white fiber data application, data depth and functional data) and aims to understand if it is possible to describe the distribution of white fibers in the brain with data depth? How to define a data depth that is invariable in the choice of curve parametrization and thus to produce a model of brain white fibre trajectories?

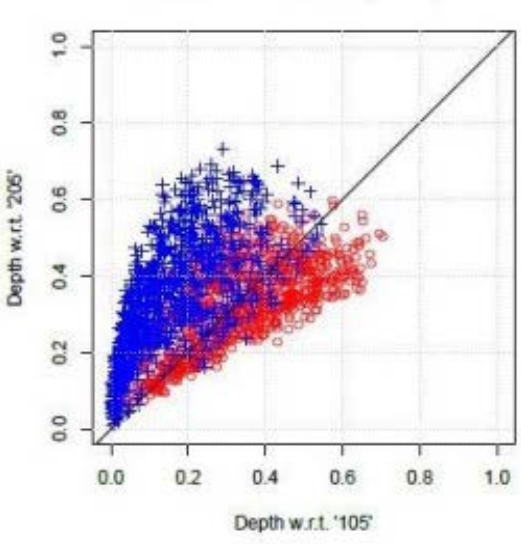
Neuronal fibers, also called axons, are extensions of nerve cells that transmit electrical information between different regions of the brain. The depth of data was originally introduced in a seminal paper by Tukey (1975). Since then, several other notions of depth of data have been proposed.
As a result of these theoretical developments, it has become possible to extend standard rank-based univariate descriptive statistics to analyse multivariate observations (see, for example, Oja 1983; Liu et al. 1999). These new techniques have proved very useful for data visualization, for estimating a measure of location or propagation, for detecting outliers (see also Hubert et al. 2015), for clustering, or for detecting whether two groups of functions come from the same population. However, the functional depths are sensitive to the parameterization of the curves. According to the research of Lafaye de Micheaux, Mozharovky and Vimond, a mathematical curve describing a given fiber must be understood as the set of all points that describe the spatial location of one of these fibers, without any focus on any parameterization. Indeed, as Kurtek and his collaborators (2012) point out, «parameterization is simply intended to facilitate analysis and is not an intrinsic property of a curve», which leads them to recommend that «shape analysis should be invariant not only for rigid movements and global scaling, but also for their parameterizations». With this motivation in mind, the scientists developed a new concept of depth for curves that is invariant to the choice of parameterization. It will be widely applicable, thanks to the freely available R/C++ package curveDepth (Mozharovskyi et al. 2019), to many other similar types of data.
The paper details the research of Lafaye de Micheaux, Mozharovky and Vimond in a series of sessions that first aim at formally defining the curves and introducing a model for the sampled curves. Then to define a new data depth for the curves. After addressing implementation issues, they present the results of the simulations. They also apply the depth of their curves to analyze brain imaging datasets and to classify handwritten numbers. The depth of their curve is compared to other existing depths, namely the MFHD (Claeskens et al. 2014), the modified multivariate band depth (mMBD) of Ieva & Paganoni (2013), the skewed multivariate functional bias adjusted projection depth (saPRJ) of Hubert et al. (2015), the simplistic band depth (SBD) of López-Pintado et al. (2014) and its modified version (mSBD).

Finally, Lafaye de Micheaux, Mozharovky and Vimond collect some concluding remarks. It gathers all the technical evidence, as well as the codes and data necessary to reproduce all their numerical and graphical results. These results suggest that the calculations converge with the hypotheses developed by the scientists. The next step in the research will be to conduct the large-scale study based on data from several thousand pairs of twin brains.

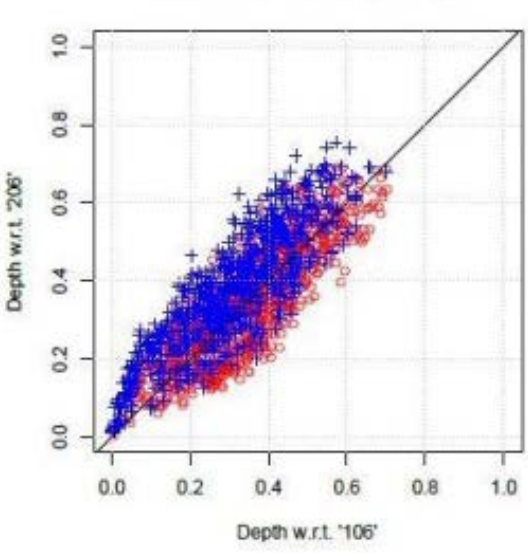
The autors thanks professor Wei Wen, from UNSW Sydney for providings the OATS data.



Illustrations of the ordering of the white matter fibers for one subject using yhe curve depth. Whole brain fiber data set for one twin
See http://biostatisticien.eu/DataDepthFig8 for an interactive 3D applet.



Depth of 105 (red) and 205 (blue)



Depth of 106 (red) and 206 (blue)

References

Original title of the article: Depth for curve data and applications
Published in: Journal of the American Statistical Association, Published online: 22 Apr 2020
Available at: https://www.tandfonline.com/doi/full/10.1080/01621459.2020.1745815

CREST
CENTER FOR RESEARCH
IN ECONOMICS AND STATISTICS

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

The Center for Research in Economics and Statistics (CREST) is a leading French scientific institution for advanced research on quantitative methods applied to the social sciences.

CREST is a joint interdisciplinary unit of research and faculty members of CNRS, ENSAE Paris, ENSAI and the Economics Department of Ecole Polytechnique. Its activities are located physically in the ENSAE Paris building on the Palaiseau campus of Institut Polytechnique de Paris and secondarily on the Ker-Lann campus of ENSAI Rennes.

cnrs

GENES