Depth for curve data and applications

Pavlo Mozharovskyi

Abstract

John W. Tukey (1975) defined statistical data depth as a function that determines centrality of an arbitrary point with respect to a data cloud or to a probability measure. During the last decades, this seminal idea of data depth evolved into a powerful tool proving to be useful in various fields of science. Recently, extending the notion of data depth to the functional setting attracted a lot of attention among theoretical and applied statisticians. We go further and suggest a notion of data depth suitable for data represented as curves, or trajectories, which is independent of the parametrization. We show that our curve depth satisfies theoretical requirements of general depth functions that are meaningful for trajectories. We apply our methodology to diffusion tensor brain images and also to pattern recognition of hand written digits and letters. Supplementary Materials are available online (see https://arxiv.org/abs/1901.00180).