# Accepted Manuscript


Connecting pairwise geodesic spheres by depth: DCOPS

Ricardo Fraiman, Fabrice Gamboa, Leonardo Moreno

PII: S0047-259X(17)30734-0
DOI: https://doi.org/10.1016/j.jmva.2018.08.008
Reference: YJMVA 4403

To appear in: *Journal of Multivariate Analysis*

Received date : 29 November 2017

Please cite this article as: R. Fraiman, F. Gamboa, L. Moreno, Connecting pairwise geodesic spheres by depth: DCOPS, *Journal of Multivariate Analysis* (2018), https://doi.org/10.1016/j.jmva.2018.08.008

# Connecting pairwise geodesic spheres by depth: DCOPS

Ricardo Fraiman [‡a], Fabrice Gamboa[†b], Leonardo Moreno[*c]

[a‡]Centro de Matemática, Facultad de Ciencias (UdelaR), and Instituto Pasteur de Montevideo, Uruguay
[b†]Institut de Mathématiques de Toulouse, France
[c*]Departamento de Métodos Cuantitativos, FCEA, Universidad de la República, Uruguay

**Abstract**

We extend the classical notion of spherical depth for data in $\mathbb{R}^k$ to the case of data on a Riemannian manifold. We show that this notion has several desirable properties. The uniform consistency and limiting distribution of an empirical analogue of this depth function are studied. Consistency is also shown for functional data. Various illustrations involving Riemannian manifold data are included.

*Keywords:* Depth measures, geodesic distance, Riemannian manifolds, spherical depth.

## 1. Introduction

Depth concepts have become a fundamental tool in modern statistics. Over the last three decades, data depth has emerged as a powerful notion which leads to a center-outward ordering of multivariate data and, more recently, functional data. Beyond leading to the notions of robust multivariate center (median) and trimmed means, the concept has found many other applications in recent years. It has been applied successfully to supervised and unsupervised learning, hypothesis testing, and outliers detection; see, e.g., the seminal paper by Liu et al. [27].

The notion of depth was first introduced in [48] for bivariate data sets and it was extended to higher dimensions in [14]. Several other notions were later introduced, such as convex hull peeling depth [4], Oja depth [35], simplicial depth [25], spatial depth [49] and spherical depth [17], among others. In particular, several notions have been introduced in the functional data setting in the recent past.

Half-space, simplicial, spatial depth and spherical depth are fundamental and their properties have been studied extensively. They share a list of desirable properties that were introduced in [25] for the simplicial depth and studied in [44] for general depth notions. These desirable properties are the following.

P1) Invariance with respect to a group of transformations. Affine invariance or orthogonal invariance.

P11: Affine invariance: The depth should not depend on the coordinate system and, in particular, on the scales of the underlying measurements.

P12: Orthogonal invariance: The depth should not depend on the coordinate system and the global scale.

P2) Maximality at center: For a distribution having a uniquely defined "center", such as the point of symmetry with respect to some notion of symmetry (invariance under a suitable family of transformations like angular, central, elliptical or spherical symmetry), the depth should attain a maximum value at this center. Observe that all of these notions of symmetry coincide with the usual notion of symmetry for one dimensional data.

P3) Monotonicity relative to deepest point: When a point moves away from the "deepest point", through a ray, the depth decreases.

P4) Vanishing at infinity.

Serfling [43] also mentioned some key perspectives that are relevant to the choice of a procedure in nonparametric multivariate analysis. In particular to take into account dimensionality ("Typically, data in $\mathbb{R}^k$ has structure of lesser dimension than the nominal $k$") and computional complexity ("The feasibility of computation is a function of the size $n$ and and dimension $k$ of the data, with practical limitations arising in the case of higher $k$.")