Summary: Big data, high dimensional data, sparse data, large scale data, and imaging data are all becoming new frontiers of statistics. Changing technologies have created this flood and have led to a real hunger for new modeling strategies and data analysis by scientists. In many cases data are not Euclidean; for example, in molecular biology, the data sit on manifolds. Even in a simple non-Euclidean manifold (circle), to summarize angles by the arithmetic average cannot make sense and so more care is needed. Thus non-Euclidean settings throw up many major challenges, both mathematical and statistical. This paper will focus on the PCA and clustering methods for some manifolds. Of course, the PCA and clustering methods in multivariate analysis are one of the core topics.

We basically deal with two key manifolds from a practical point of view, namely spheres and tori. It is well known that dimension reduction on non-Euclidean manifolds with PCA-like methods has been a challenging task for quite some time but recently there has been some breakthrough. One of them is the idea of nested spheres and another is transforming a torus into a sphere effectively and subsequently use the technology of nested spheres PCA. We also provide a new method of clustering for multivariate analysis which has a fundamental property required for molecular biology that penalizes wrong assignments to avoid chemically no go areas. We give various examples to illustrate these methods. One of the important examples includes dealing with COVID-19 data.

MSC:

62H11 Directional data; spatial statistics
62H15 Hypothesis testing in multivariate analysis
62P10 Applications of statistics to biology and medical sciences; meta analysis

Keywords:

adaptive linkage clustering; circular mode hunting; dimension reduction; multivariate wrapped normal; SARS-CoV-2 geometry; stratified spheres; torus PCA

Full Text: DOI

References:


[18] Huckemann, S. F., Comments on: Recent advances in directional statistics, TEST, 1-5 (2021) · Zbl 1474.62465


[45] Sommer, S., Horizontal dimensionality reduction and iterated frame bundle development, (Geometric Science of Information (2013), Springer), 76-83 · Zbl 1356.62012


[49] Yang, C.-H.; Vemuri, B. C., Nested grassmanns for dimensionality reduction with applications to shape analysis, (International


This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.