Research Highlights

Article

Junyoung Park, Changwon Yoon, Cheolwoo Park, and Jeongyoun Ahn
Kernel Methods for Radial Transformed Compositional Data with Many Zeros,

Contents

Compositional data analysis with a high proportion of zeros has gained increasing popularity, especially in chemometrics and human gut microbiomes research. Statistical analyses of this type of data are typically carried out via a log-ratio transformation after replacing zeros with small positive values. We should note, however, that this procedure is geometrically improper, as it causes anomalous distortions through the transformation. We propose a radial transformation that does not require zero substitutions and more importantly results in essential equivalence between domains before and after the transformation. We show that a rich class of kernels on hyperspheres can successfully define a kernel embedding for compositional data based on this equivalence. The applicability of the proposed approach is demonstrated with kernel principal component analysis.

References

Aitchison, J. Principles of compositional data analysis. Lecture Notes-Monograph Series, pp. 73–81, 1994.

Muandet, K., Fukumizu, K., Sriperumbudur, B., and Sch¨olkopf, B. Kernel mean embedding of distributions: A review and beyond. Foundations and Trends in Machine Learning, 10(1-2):1–144, 2017.