THE UNIT SIMPLEX -- SAMPLE SPACE FOR COMPOSITIONAL DATA.

1Bren, M., 2Martín Fernández, J.A., 1University of Maribor, FOV, Kranj, Slovenia, 2University of Girona, Dep. for Inf. and App. Math., Girona, Spain

It is well known that the adequate sample space for compositional data is the unit simplex - which , equipped with the perturbation operation and scalar power multiplication, is a real vector space. We'll discuss the centered logratio transformation, a homomrphism of the n-dimensional unit simplex onto a n-1 dimensional real hyperplane - subspace of the usual Euclidean real vector space. We'll define and also discuss the inner product in the vector space of the unit simplex and the centered logratio transformation as a homeomorphism of the unitary spaces that preserves oientation, angles between vectors, and transforms lines into lines. Examples of simulated data sets and of real geological data sets will be presented to visualize the three way compositions in ternary diagrams and in a real plane with the usual Euclidean metric which is a benefit because of our Euclidean seeing. We also stress that all statistical techniques: principal components, discriminant analysis, cluster analysis... are grounded on this Euclidean seeing.