Learning Dual contextual/conceptual geometries of Databases/Matrices

Ronald Coifman (Yale University)

We provide an overview of recent developments in geometric Harmonic Analysis methodologies for empirical organization of stochastic data. We focus on data provided as an array or matrix, where we view the rows and columns of the matrix as being in a functional duality, generating joint row/column organizational geometries, and opening the door to automated analytic organizations of Matrices or Databases.

We show that the various metrics generated can be viewed as generalizing Earth mover metrics, in which geometry and statistics are mutually supportive. In particular, we introduce methodologies extending Harmonic analysis and "signal processing" on data matrices, enabling functional regression, prediction, denoising, compression, fast numerics, and so on. We illustrate these ideas to organize and map out in an automatic and purely data driven fashion, text documents, psychological questionnaires, medical profiles, physical sensor data, financial data.

As an application to Mathematics , we show that this organization enables the constructions of a dual phase space geometry of eigenfunctions of Laplace Beltrami operators (where the eigenvectors are the rows (questions) and the points are the columns), in this dual geometry we have a corresponding Heisenberg principle.