



Department of Decision Sciences

Statistics Seminars

A Generalized Approach to Principal Component Analysis

Martin Schlather

University of Mannheim

Thursday, 18th November 2021
12:00 pm Room AS02 Via Roentgen 1 Milano

Abstract

Principal Component Analysis is a well-known procedure to reduce intrinsic complexity of a standardized data set, essentially through simplifying the correlation structure. We provide a substantial extension based on semi-groups, which includes distributions without second moments. We reformulate the PCA as a best low rank approximation defined through a regression problem, which is closely connected to autoencoders and admits solutions under mild assumptions. As a specific example, we apply our general formulation to extreme value distributions. In this case solving the regression problem reduces to a tractable and practically solvable problem for real datasets. As a side effect of our general approach, the perspective on uncorrelated variables is inverted: the assertion that the covariance is zero is not the definition, but a corollary; the fact that uncorrelated variables need not be independent is not a warning, but intrinsic to the definition.