Projection pursuit via kernel mean embeddings

Lutz Duembgen

University of Bern, Switzerland

Abstract

Detecting and visualizing interesting structures in high-dimensional data is a ubiquitous challenge. If one aims for linear projections onto low-dimensional spaces, a well-known problematic phenomenon is the Diaconis–Freedman effect: under mild conditions, most projections do not reveal interesting structures but look like scale mixtures of spherically symmetric Gaussian distributions. We present a method which combines global search strategies and local projection pursuit via maximizing the maximum mean discrepancy (MMD) between the empirical distribution of the projected data and a data-driven Gaussian mixture distribution. Here, MMD is based on kernel mean embeddings with Gaussian kernels.