Restricted Kernel Canonical Correlation Aanalysis

Nina Klenovšek

Kernel canonical correlation analysis is a method analyzing the dependence between two sets of random variables when the classical method, canonical correlation analysis, fails because of the nonlinearity in data. This kernel method is mostly used in machine learning, especially for information retrieval and text mining where the data is usually represented with non-negative numbers. This gives us an idea to incorporate the non-negativity restriction directly into the problem of kernel canonical correlation analysis and propose a new method which we call *restricted kernel canonical correlation analysis* (RKCCA). With Karush-Kuhn-Tucker theory and the idea of subvectors and submatrices we translate the problem of searching for the kernel canonical correlation under non-negativity restriction into a generalized eigenvalue problem with a real symmetric matrix and a positive definite matrix. Using a simple search process we find the optimal eigenvalue which is equal to the kernel canonical correlation under non-negativity restriction we are looking for.