

Robust Dimension Reduction Based on CANCOR

Jianhui Zhou
Department of Statistics
University of Illinois at Urbana-Champaign

Abstract

The canonical correlation (CANCOR) method for dimension reduction in a regression setting is based on the first and second moments of the data, and therefore sensitive to outliers. We study a weighted canonical correlation (WCANCOR) method where each observation is weighted based on its Mahalanobis distance to the center of the predictors. Robust estimates of the location and scatter such as the minimum covariance determinant (MCD) estimator of Rousseeuw (1984) can be used to compute the Mahalanobis distance. To determine the number of significant dimensions in the WCANCOR method, we consider a weighted permutation test. A comparison between the weighted CANCOR and the weighted SIR methods is also made to see the relative merits of each method.

Key words: canonical correlation, dimension reduction, MCD estimator, permutation test, robustness.