

Low-rank approximations and weighted low-rank approximations

Paulo Canas Rodrigues

Nova University of Lisbon, Portugal

Abstract

Principal component analysis (PCA) is one of the most widely used multivariate techniques. It is usually applied to two-way matrices with individuals in the rows and variables in the columns, and converts the possibly correlated variables into a set of orthogonal variables, the principal components. Several algorithms have been proposed to obtain the least squares estimates for the component scores and for the loadings, being the most used the eigenvalue decomposition of the covariance (or correlation) matrix of the data or the singular value decomposition of the two-way data matrix. In this paper we will be mostly interested in the weighted version of this low-rank approximation. This allows us to give weights to the variables and/or the individuals according to the outcome of a preliminary analysis of the two-way data, e.g., in the case of repeated measurements the weights can be given by the inverse of their error variances. The use of the weighted PCA also increases the robustness when compared with the standard PCA. Applications to genetic and financial data will be presented.

Keywords

Principal component analysis, Additive main effects and multiplicative interaction model, Plant genetics, Public debt.