

Common components to construct biplots for three mode data

Darryn Williams* and Sugnet Lubbe

University of Cape Town, Cape Town, South Africa darryn.williams@gmail.com

Statistics is well acquainted with the use of exploratory techniques to reveal detail about the structure of the data. Graphical displays of data are key in this process, chief amongst them the scatterplot. However, a scatterplot is restricted to use with data comprising at most three variables. For higher dimensional data comprising $p > 3$ variables, the analogue to this display is the biplot. It is a plot that displays axes for all the variables that have been measured together with points to represent all the subjects in a low dimensional approximation of the p dimensional space. It is useful in that it affords the means to instantly assess correlations between variables as well as any groupings of the subjects that might be inherent in the dataset. This display is well developed for data comprising n samples and p variables. Such datasets are referred to as two mode data where the objects and variables represent modes. This paper is focused on using biplots for three mode data. More specifically, longitudinal data is considered. Generally this type of data comprises a number of objects on which a number of measurements have been made at different occasions. The modes are thus subjects, variables and occasions. Canonical Variate Analysis biplots provide a means to visually appraise the structure of grouped data. Here details are given on the construction of a Canonical Variate Analysis biplot based on Common Principal Components to allow grouped longitudinal data to be examined in a single plot. Simulated data is used for illustration.

Key Words: Common Principal Components, Canonical Variate Analysis biplots