

**Chouakria, A.; Gazes, P.; Diday, E.**

**Symbolic principal component analysis.** (English) Zbl 0977.62063

Bock, Hans-Hermann (ed.) et al., Analysis of symbolic data. Exploratory methods for extracting statistical information from complex data. Berlin: Springer. Studies in Classification, Data Analysis, and Knowledge Organization. 200-212 (2000).

Principal components analysis (PCA) takes as input a data matrix of the type  $X = (x_{kj})_{n \times p}$ , where  $(x_{kj})$  is the precise and single value of the descriptive feature  $Y_j$  for the object  $k \in \Omega = \{1, \dots, n\}$ . However, in practice the investigated objects are often more complex and so more complex data are required in order to provide an accurate description of these objects. These data are called symbolic data which finally result in symbolic objects of various types.

In this paper the focus is on one particular type of symbolic data, i.e., interval data. In this case the elements of the data matrix are intervals  $\xi_{kj} = [x_{kj}, \overline{x_{kj}}] \in \mathbb{R}$ . A new method, called Vertices Method, is proposed which extends the classical PCA to interval data. For the case of a high number  $p$  of features of interval type another method, called Centers Method, is proposed. Both methods are compared and illustrated by an example.

The proposed methods achieve two aims: on the one hand, they analyze, as the classical PCA method does, the main trend of the objects' spatial distribution and, on the other hand, they estimate and visualize for each object the resulting variability or inaccuracy.

For the entire collection see [[Zbl 1039.62501](#)].

Reviewer: [M.P.Moklyachuk \(Kyiv\)](#)

**MSC:**

[62H25](#) Factor analysis and principal components; correspondence analysis  
[62-07](#) Data analysis (statistics) (MSC2010)

Cited in **8** Documents

**Keywords:**

[symbolic objects](#); [symbolic data](#); [principal components analysis](#)