

Dependence and interdependence analysis for interval-valued variables

Carlo Lauro¹ and Federica Gioia²

¹ Dipartimento di Matematica e Statistica
Università di Napoli Federico II
Complesso Universitario di Monte S. Angelo, Via Cinthia
I-80126 Napoli, Italy

² Dipartimento di Matematica e Statistica
Università di Napoli Federico II
Complesso Universitario di Monte S. Angelo, Via Cinthia
I-80126 Napoli, Italy

Abstract. Data analysis is often affected by different types of errors as: measurement errors, computation errors, imprecision related to the method adopted for estimating the data. The methods which have been proposed for treating errors in the data, may also be applied to different kinds of data that in real life are of interval type. The uncertainty in the data, which is strictly connected to the above errors, may be treated by considering, rather than a single value for each data, the interval of values in which it may fall: *the interval data*. The purpose of the present paper is to introduce methods for analyzing the *interdependence* and *dependence* among *interval-valued* variables. Statistical units described by interval-valued variables can be assumed as a special case of Symbolic Object (SO). In Symbolic Data Analysis (SDA), these data are represented as boxes. Accordingly, the purpose of the present work is the extension of the *Principal Component Analysis* to obtain a visualization of such boxes, on a lower dimensional space. Furthermore, a new method for fitting an *interval simple linear regression* equation is developed. With difference to other approaches proposed in the literature that work on scalar recoding of the intervals using classical tools of analysis, we make extensively use of the interval algebra tools combined with some optimization techniques.

Keywords

INTERVAL-VALUED VARIABLE, INTERVAL ALGEBRA, INTERVAL CORRELATION MATRIX, INTERVAL EIGENVECTORS, INTERVAL EIGENVALUES, INTERVAL REGRESSION LINE, VISUALIZATION