The Path Point Approach to Dimension Reduction and Visualization of Multiple Time Series Data

吳漢銘
臺北大學統計學系

Abstract

Exploratory analysis and visualization, which are performed prior to the modeling and forecasting of multiple time series data, present important steps in the discovery of the underlying dynamics of a series. In this study, we present extensions of two dimension reduction (DR) methods, principal component analysis (PCA) and sliced inverse regression (SIR), to multiple time series data through a symbolic data analysis approach, named the path point approach. First, the multiple time series data are put in the form of time-dependent intervals such that the interval is described by a starting value and an ending value of a certain time period. In this context, each series can be geometrically represented as consecutive directed segments. Then, we apply PCA and SIR to the starting and ending points of those directed segments to visualize the temporal trajectories of objects in a lower dimensional subspace. Experimental results from studies of microarray time series data concerning a yeast cell cycle and the financial data show that the proposed methods are capable of providing insight into the structure and behaviors of objects on the 2D factorial axes. Comparisons with the currently existing method, namely, the applied vector approach, for PCA and SIR for time-dependent interval data are also examined.

Keywords: exploratory data analysis, data visualization; PCA; sliced inverse regression; symbolic data analysis; time dependent interval-valued data.