Dimension reduction for mixed-type data

陳定立[1], 須上英[2]

[1] 中央研究院, [2] 台北大學

Abstract

Dimension reduction is often an important data pre-processing procedure before further analysis. In this talk, we will present a new approach for dimension reduction. We propose to assign the data in a lower dimensional space so that some pre-defined measurement error is minimized. It can be shown that this approach is equivalent to Principal Component Analysis (PCA), one of the most popular dimension reduction methods, for continuous type of data, while this approach can be directly generalized to dimension reduction for categorical data as well as mixed-type data. A simulation study will be given to demonstrate the result of this dimension reduction method.

Keyword: Dimension reduction, pricipal component analysis, caetegorical data