SCV 2008: December 1-3, Taipei, Taiwan (Abstract)

Influence analysis in ordinary, kernel and functional multivariate methods

Yutaka Tanaka
Nanzan University, Nagoya, Japan

Abstract

In recent years there have been developed kernel and functional multivariate methods as extensions to ordinary multivariate methods to analyze nonlinear relations and to deal with continuous functions instead of multidimensional observations. In this presentation we discuss influence or sensitivity analysis in these three types of multivariate methods.

The basic idea is to derive influence functions or differential coefficients in perturbation parameters and approximate the influence of a subset of observations as the sum of the influences of the observations which belong to the subset on the basis of the additive property of the influence (Tanaka, 1994), that is, the generalized influence function for a subset is defined as a function of the sum of the influence functions for observations in the subset. Therefore, we may search for subsets of the observations which have large norms of the influence functions to detect influential subsets of observations, where the inverse of the covariance matrix of the estimator for parameter vector is used as the metric to compute the norm. We have developed a software SAMMIF for sensitivity analysis in ordinary multivariate methods (Mori et al., 1998).

The same technique can be used to develop methods of influence analysis in kernel multivariate methods, say, in kernel principal component analysis. In kernel methods observations are mapped to a higher dimensional space. Our idea is to approximate the higher dimensional space by lower dimensional space by using principal component analysis (PCA) or other dimension reduction methods and then utilize the influence analysis in the corresponding ordinary multivariate method. We will show an example of kernel PCA of the gene network data (Tanaka and Yamanishi, 2007).

Methods of influence analysis can be developed in functional data analysis, e.g., functional PCA (Yamanishi and Tanaka, 2005). In functional data analysis a smoothing parameter lambda is introduced in estimating weight functions, and the lambda is estimated using cross-validation. We derived influence functions for weight functions of functional PCA with and without taking into account the influence through the lambda.

Influence of variables on the results of analysis will also be discussed.

References.

Mori,Y. et al.(1998). COMPSTAT 98 (Eds. R.Pain and P.Green) Physica Verlag, 395-400.

Tanaka,Y. (1994). J. Japanese Soc. Comp. Statist., 2, 9-20.

Tanaka,Y. and Yamanishi,Y. (2005). IASC-ARS Special Conference, Seoul, Korea.

Yamanishi,Y. and Tanaka,Y.(2007). Computaional Statistics, 20, 311-326.