THE SUMMARY OF Ph.D. DISSERTATION

Major		SURNAME, Firstname TAKEUCHI, Hidekazu

Title

Assessment of Influence of Observations in Linear Regression

Abstract

In this dissertation we investigate assessment of influence out of major problems in linear regression. We mainly focus on the assessment of influence of observations on various influence measures based on estimates, predicted values, residuals and so on in linear regression. We also take up the assessment of influence of parameters included in estimation procedures on the behavior of various measures.

Although it is needed to apply some influence assessment procedures in practical data analysis, there still remain various problems to be addressed from both theoretical and practical points of view for their sound applications. Hence to overcome them we here propose some practical procedures, give concrete guidelines for the assessment of influence and evaluate them.

Chapter 1 is an introduction to the assessment of influence of observations, and gives the purpose and the organization of this dissertation.

In Chapter 2 we propose new expressions of influence measures to assess the influence of observations in least squares estimation. Especially the new expression of Cook's distance in the multiple case where we detect jointly influential observations is based on the expression in the single case where we detect individually influential ones and it gives us an effective method to detect the masking effect.

In Chapter 3 new expressions of influence measures are given to assess the influence in ridge estimation in place of least squares one. The new expressions enable us to investigate basic causes of the influential observations separately. Furthermore the behavior of the ridge parameter is investigated through auxiliary influence indices derived from the new expressions.

A new influence assessment procedure is given in variable selection problems in Chapter 4. We propose a new influence measure to assess the influence of individual observations on prediction mean square errors from a viewpoint different from the existing ones. The effectiveness of the new influence measure is shown by numerical examples. Furthermore it is shown that the new influence measure is expressed in terms of Mallows' Cp statistic and Cook's distance.

In Chapter 5 we investigate the behavior of the power function when we use Cook's distance to detect outliers in the sense of residuals. It is shown that Cook's distance is quite unstable in its power through its application to real data sets.

In Chapter 6 we propose a class of generalized inverse regression estimators in multi-univariate linear calibration by combining the classical and the inverse regression estimators. We derive a pertinent estimator which improves both of the two, and show some properties of mean square errors of the proposed estimator in terms of the parameter.

Finally we summarize this study on the assessment of influence of observations in linear regression in Chapter 7.