

Nonparametric Principal Components Regression

by

Jennifer D. Umali

A thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science (Statistics)

School of Statistics
University of the Philippines
Diliman, Quezon City

March 2010

ABSTRACT

The issue of specification bias and multicollinearity are often encountered when modeling in a high dimensional data. Principal components regression resolves the multicollinearity problem but not the specification bias due to the selection of the important principal components to be included in the model, resulting to the deterioration of predictive ability of the model. The problem on the deterioration of the model fit while resolving the issue on multicollinearity brought about by the dimensionality problem can be addressed by formulating the principal components regression in a nonparametric framework.

The regression model is proposed to be estimated by first extracting the principal components of the independent variables then a nonparametric regression on principal components is fitted. Using simulated data, the proposed nonparametric principal components regression is shown to address the multicollinearity problem and resolve the issue of high dimensionality while retaining higher predictive performance relative to the parametric counterpart.

Keywords: High dimensional data, multicollinearity, principal components regression, nonparametric regression