# Inference in Linear Multivariate Envelope Model

Tat Yau, MSc

Louisiana State University Health Sciences Center, Biostatistics Program
New Orleans, LA

## ABSTRACT

For the past decade, big data have become more accessible with the growth of sophisticated technology. Various methodologies are thus developed to handle big data. Envelope model, developed in Cook et al. (2010), is a parsimonious version of the classical multivariate linear regression model with the intention to achieve dimension reduction. The developed response envelope model particularly addresses the possibility of dimension reduction in response variables, by considering that some linear combinations of the response variables are not related to the independent variables. With the restriction that the subspace spanned by the columns of the matrix parameter of expectation is contained in the column space of the covariance matrix of the error in envelop model, efficiency is gained for estimating the matrix parameter. Compared to the ordinary least square (OLS) estimators, the maximum likelihood estimators (MLEs) of the matrix parameters based on the envelope model are asymptotically less variable. This variance reduction is achieved by projecting the OLS estimators onto the proposed envelope subspace, for which its dimension is assumed to be fixed throughout the derivation of envelope model. In application, certain likelihood based criteria are suggested to estimate the dimension of the envelope subspace prior to estimating the model parameters. The primary interest of this research is the inference in envelope model, specifically, hypothesis testing about the mean parameters and the statistical properties of the envelope estimator. As the large sample statistical properties are derived under the assumption that the envelope dimension is fixed, the derived asymptotic covariance matrix of the mean parameter is in fact ignoring the variability of estimating the envelope subspace dimension. In this dissertation, we focus mainly on two aspects: first, for a finite sample size, how well the selected criteria perform for selecting the correct envelope dimension, if the "true" envelope model is among investigated models; secondly, we investigate the finite sample properties of the Wald test under envelope model, in which the envelope estimator and its large sample covariance matrix are implemented. A simulation study has been designed for this two-fold purpose. The surrogate for the distribution of the test statistic under the null, the distribution of p-value, is studied.