BOOTSTRAPPING FOR A STRUCTURAL EQUATION MODEL WITH A NEARLY NON-POSITIVE DEFINITE FITTED COVARAINCE MATRIX

YUSEP SUPARMAN

Abstract. Structural equation modeling (SEM) has been widely adopted for measuring causal relationship. Instead of fitting individual observations as in regression analysis, SEM works in a different way by fitting a sample covariance matrix to a model implied covariance matrix and producing a fitted covariance matrix. One problem emerges when a fitted covariance matrix is nearly non-positive definite. The estimated maximum likelihood standard errors in the model become implausibly large which in turn jeopardize inferences. In this paper, we show that bootstrapping can overcome this problem. We compare standard errors from two equivalent SEM models, i.e. a discrete and a continuous time hedonic price autoregression panel model. We find out that the bootstrap standard errors obtained from the continuous time model, which suffers from nearly non-positive definite fitted covariance matrix problem, are comparable to the maximum likelihood standard errors obtained from the discrete time model which is free from the problem.

Key words: autoregressive panel model, bootstrap, exact discrete time model, fitted covariance matrix, hedonic price model, structural equation modeling.

1. INTRODUCTION

Structural equation modeling (SEM) has been widely adopted in behavior and life sciences for examining causal relationships. One of the reasons is that SEM accommodates a wide range of statistical models, from simple linear regression model to continuous time model with latent variables. In addition, SEM ability in including measurement errors into the analysis overcomes methodological obstacle which cannot be handle in a standard statistical analysis. Furthermore, SEM allows latent constructs to be explicitly included in the model which enhance the interpretations of analysis results.

Different from regression analysis which works based on individual data, SEM works based on a covariance matrix. SEM fits a sample covariance matrix to a model implied covariance matrix which is a function of parameters in the model and produces a fitted covariance matrix (Jöreskog, 1996). Among estimation methods, the maximum likelihood has been the most often used. Nevertheless, one problem arise when a fitted covariance matrix is non-positive definite or nearly non-positive definite. In the former condition, we could not estimate the standard error. In the later one, we obtain implausibly large standard error. The two conditions jeopardize the parameter estimates inference. This problem emerges due to the maximum likelihood standard error estimates are a function of the inverse of fitted covariance matrix (Jöreskog, 1973),