Dissertation Title: Linear and Nonlinear Estimation with Spatial Data

Chapter 1  Pseudo Generalized Least Squares Regression Estimation with Spatial Data

Abstract  It is hard to account for all pairwise correlations in the estimation of the mean parameters for a large sample of spatial data. In a linear regression model, a pseudo generalized least squares (PGLS) approach is proposed. Data could be divided into groups according to natural geographic areas, only correlations within groups are accounted for. Since correlations within groups account for most of the correlations among observations, the resulting PGLS estimator will not lose much efficiency compared to GLS. The PGLS estimator is consistent and asymptotically normal, and computationally easier than GLS. A HAC variance covariance estimator which is robust to both heteroskedasticity and spatial correlation is provided. Moreover, a convenient way of using the OLS residuals to calculate the spatial correlation parameters is also proposed, which is very easy and works well as shown in the simulation. Simulation study also shows that PGLS is more efficient than OLS under certain conditions. This paper also gives an empirical example on estimating policy effects on student performance accounting for the correlations among schools within the same school districts.


Abstract  Generalized least squares can be used to improve the estimation efficiency in a linear regression model even if the variance covariance structure is misspecified. Similarly, generalized estimating equations (GEE) or weighted multivariate nonlinear least squares (WMNLS) are used in nonlinear panel data models and system of equations to obtain a more efficient conditional mean parameters. In this paper, GEE is applied to cross section data with spatial correlations in nonlinear models. I use a pooled quasi-maximum likelihood estimator (PQMLE) in the first step and use GEE approach in the second step. Given some regularity conditions and assumptions, the asymptotic distribution of the two-step estimator is derived in the framework of M-estimation. I use a probit model with a latent spatially correlated error term and a count data model with a multiplicative spatial error term to demonstrate the GEE procedures. Monte Carlo simulation shows efficiency comparison of the PQMLE and GEE. The results show correct modeling the structure of the working correlation matrix is quite important in nonlinear models, which is quite different from the linear model.

Chapter 3  Conditions for the Numerical Equality of the OLS, GLS and Amemiya-Cragg Estimators

Abstract  Conditions under which the ordinary least squares (OLS) and generalized least squares (GLS) estimators are equal are well known. This paper extends these results in two ways. First, we give conditions under which GLS based on one assumed error variance matrix equals GLS based on a different assumed variance matrix. Second, we give conditions under which GLS equals the GMM estimator of Amemiya (1983) and Cragg (1983).