Feb 14, 2019

1 Spatial Autoregressive (SAR) Model

1.1 SAR process

$$y_i = \lambda_{i,n} Y_n + \epsilon_i,\ i = 1, \dots, n$$

where $Y_n = (y_1, \dots, y_n)'$ is the column vector of dependent variables, $w_{i,n}$ is a n-dimensional row vector of constants, and ϵ_i's are i.i.d. $(0, \sigma^2)$. In matrix form,

$$Y_n = \lambda W_n Y_n + \mathcal{E}_n$$

$W_n Y_n$ called 'spatial lag' and under assumption $S_n(\lambda) = I_n - \lambda W_n$

$$Y_n = S_n^{-1}(\lambda)\mathcal{E}_n$$

the regression model with SAR disturbance U_n is specified as

$$Y_n = X_n\beta + U_n,\ U_n = \rho W_n U_n + \mathcal{E}_n$$

where $\mathcal{E}_n$ has zero mean and variance $\sigma^2 I_n$

1.2 Mixed regressive, spatial autoregressive model (MRSAR)

$$Y_n = \lambda W_n Y_n + X_n\beta + \mathcal{E}_n$$
$$Y_n = S_n^{-1}(\lambda) X_n\beta + S_n^{-1}(\lambda)\mathcal{E}_n$$

where $\mathcal{E}_n$ has zero mean and variance $\sigma^2 I_n$

1.3 Other models

A more rich SAR model may combine the MRSAR equation with SAR disturbance

$$Y_n = \lambda W_n Y_n + X_n\beta + U_n,\ U_n = \rho M_n U_n + \mathcal{E}_n$$

Further extention of a SAR model may allow high-order spatial lags as in

$$Y_n = \sum_{j=1}^{p} \lambda W_{jn} Y_n + X_n\beta + \mathcal{E}_n,$$

where W_{jn}'s are p distinct spatial weights matrices.

2 Estimation Method

- Maximum Likelihood Estimator (MLE) : has usually goodfinite sample properties relative to other methods for the estimation of SAR models with the first order of spatial lag. However the ML method is not computationally attractive for models with more than one single spatial lag.

- 2Stage Least Square (2SLS or IV) : Applicable only for the MRSAR model. Need orthogonality condition

- Generalized Method of Moment (GMM) : With properly designed moment equations, the best GMM estimator exists and can be asymtotically efficient as th ML estimate under normal disturbances

2.1 ML for SAR process

Under assumption that $\mathcal{E}_n$ is $N(0, \sigma^2 I_n)$.

$$\ln L_n(\lambda, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2} ln\sigma^2 \ln|S_n(\lambda)| - \frac{1}{2\sigma^2} Y_n' S_n'(\lambda) S_n(\lambda) Y_n$$

where $S_n(\lambda) = I_n \lambda W_n$

For the regression model with SAR disturbances, the log likelihood function is

$$\ln L_n(\lambda, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 \ln|S_n(\lambda)| - \frac{1}{2\sigma^2}(Y_n - X_n\beta)' S_n'(\lambda) S_n(\lambda)(Y_n - X_n\beta)$$

The log likelihood function for MRSAR model is

$$\ln L_n(\lambda, \sigma^2) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 \ln|S_n(\lambda)| - \frac{1}{2\sigma^2}(Y_n S_n(\lambda) - X_n\beta)'(Y_n S_n(\lambda) - X_n\beta)$$

2.2 2SLS Estimation for the MRSAR model

- MRSAR : $Y_n = \lambda W_n Y_n + X_n\beta + \mathcal{E}_n$

- the spatial lag $W_n Y_n$ can be correlated with the disturbance $\mathcal{E}$

- To avoid the bias due to correlation of $W_n Y_n$ with $\mathcal{E}_n$ need to use the Instrumental Variables(IVs)

- Let Q_n be a matrix of instrumental variables

- Denote $Z_n = (W_n Y_n, X_n)$ and $\theta = (\lambda, \beta')'$. The MRSAR equation can be rewritten as $Y_n = Z_n\theta + \mathcal{E}$

2SLS estimator of θ with Q_n is

$$\hat{\theta}_{2sls,n} = [Z_n' Q_n (Q_n' Q_n)^{-1} Q_n' Z_n]^{-1} Z_n' Q_n (Q_n' Q_n)^{-}1 Q_n' Y_n$$

Orthogonality condition is $Q_n'\mathcal{E}_n$

2.3 Method of Moments

$$\min_{\theta} g_n'(\theta) g_n(\theta)$$

where

$$g_n(\theta) = (Y_n' S_n'(\lambda) S_n(\lambda) Y_n - n\sigma^2, Y_n' S_n'(\lambda) W_n' W_n S_n(\lambda) Y_n - \sigma^2 tr(W_n' W_n), Y_n' S_n'(\lambda) W_n S_n(\lambda) Y_n)'$$

It means that find the parameter which has the minumun bias

$$E(\mathcal{E}_n'\mathcal{E}) = n\sigma^2,\ E(\mathcal{E}_n' W_n' W_n \mathcal{E}) = \sigma^2 tr(W_n', W_n),\ E(\mathcal{E}_n' W_n \mathcal{E}) = 0$$

2.4 GMM

For the MRSAR model

- Q_n be an IV matrix constructed as functions of X_n and W_n
- Let $\epsilon_n(\theta) = S_n(\lambda) Y_n - X_n\beta$, Orthogonal conditions are $Q_n'\epsilon_n(\theta)$

m is a finite number, and $P_{1n}, \ldots, P_{mn}$ are constant matrices which each has a zero diagonal.

$$g_n(\theta) = (P_{1n}\epsilon(\theta), \ldots, P_{mn}\epsilon(\theta), Q_n)'\epsilon_n(\theta)$$

3 Real Data

3.1 A Network Model with Social Interactions

$$Y_r = \lambda W_r Y_r + X_r\beta_1 + W_r X_r \beta_2 + L_{m_r}\alpha_r + u_r,\ u_r = \rho M_r u_r + \epsilon_r$$

- $\epsilon_r = (\epsilon_{1r}, \ldots, \epsilon_{m_r r})',\ \epsilon_{ir}$ i.i.d $(0, \sigma^2)$

- R is number of groups
- m_r is number of members in r group,
- W_r, M_r exogenous network (social) matrices

Lin(2005, 2006) - AddHealth Data; (https://www.cpc.unc.edu/projects/addhealth/documentation/publicdata)

- The Add Health Survey
- Student in grade 7-12; 132 schools
- Wave I in school survey - 90,118 students
- Friendship network - name up to 5 male and 5 female friends