Escuela de Estadística
Universidad Nacional de Colombia - Sede Medellín

Estimation of Social Interaction models

A Bayesian Approach

Advisor: Juan Carlos Salazar
Advisor: Andrés Ramírez Hassan

October 12, 2017

Characteristics

- New models in the branch of social economy
- Consider interdependencies among a group of individuals.
- Many economic, political and social interactions are formed by the structure of relationships.
- Examples

Characteristics

- These models were developed initially by sociologists who worked on the conceptualization of the problem.
- Then the statisticians and economists worked on the methodology.
- From the point of view of the econometric dimension, the origin could be traced back to the field of spatial statistics and econometrics.

Lee (2007) proposed the following specification for the SAR model:

$$Y_g = \lambda W_g Y_g + X_{g,1}\beta_1 + W_g X_{g,2}\beta_2 + \mathbf{1}_{m_g}\alpha_g + \epsilon_g \tag{1}$$

$$\epsilon_g \sim N_{m_g}(0, \sigma^2 I_{m_g}) \tag{2}$$

- $Y_g = (y_{g,1}, \cdots, y_{g,m_g})'$ is a vector of endogenous variables.
- X_g is the matrix of exogenous variables of dimension $m_g \times K$.
- $\mathbf{1}_{m_g}$ is a vector of ones of dimension of $m_g \times 1$.
- α_g are unobserved group specific effects.
- W_g is the social interaction matrix. Each of its entries will take value 1 if there is a link and 0 otherwise.

Structure of the model

- The observational units of the model are individuals, which are denoted by i
- These individuals are grouped into a single group denoted g
- The composition of the group is established before the statistical exercise
- The specification of the interactions of each group is representing by the matrix W_g
- The interaction of all groups could also be represented into one big matrix, which is a block diagonal matrix.

$$W_g = \frac{1}{m_g - 1}(\mathbf{1}_{m_g}\mathbf{1}'_{m_g} - I_{m_g}) \quad g = 1, \ldots, G \tag{3}$$

where $\mathbf{1}_{m_g}$ is the m_g-dimensional column vector of ones, and I_{m_g} is the m_g-dimensional identity matrix. Equivalently, in terms of each unit i from a group g, In terms of each unit i from a group g,

$$Y_g = \lambda \left(\frac{1}{m_g - 1} \sum_{j=1, j \neq i}^{m_g} y_{g,j} \right) + x'_{g,i,1} \beta_1 + \left(\frac{1}{m_g - 1} \sum_{j=1, j \neq i}^{m_g} x'_{g,j,2} \beta_2 \right) + \alpha_g + \epsilon_{g,i} \tag{4}$$

Considering again the social interaction model:

$$Y_g = \lambda W_g Y_g + X_{g,1}\beta_1 + W_r X_{g,2}\beta_2 + \mathbf{1}_{m_g}\alpha_g + \epsilon_g , g = 1, 2, \ldots G \tag{5}$$

where

$$\begin{aligned} Z_g &= (X_{g,1}, W_g X_{g,2}) \\ \beta &= (\beta_1', \beta_2')' \end{aligned}$$

This could be rewritten as:

$$Y_g = A_g^{-1}(Z_g\beta + \mathbf{1}_{m_g}\alpha_g + \epsilon_g) \tag{6}$$

Where

$$A_g(\lambda) = I_{m_g} - \lambda W_g$$

$$A_g = A_g(\lambda)$$

Now let

$$J_g = I_{m_g} - \frac{1}{m_g}\mathbf{1}_{m_g}\mathbf{1}'_{m_g}$$

- be the group mean projector.
- We consider the orthonormal matrix of J_g given by $[F_g, \mathbf{1}_{m_g}/\sqrt{m_g}]$.
- The columns in F_g are eigenvectors of J_g corresponding to the eigenvalues of one.
- Therefore $F'_g\mathbf{1}_{m_g} = 0, F'_g F_g = I_{m_g^*}$ and $F_g F'_g = J_g$ where $m_g^* = m_g - 1$

Pre-multiplication of the equation (6) by F_g' leads to the following transformed model without $\alpha_g's$

$$(F_g' A_g F_g)(F_g' Y_g) = (F_g' Z_g \beta) + (F_g' \epsilon_g)$$

$$A_g^* Y_g^* = Z_g^* \beta + \epsilon_g^*$$

where

$$A_g^* = (F_g' A_g F_g)$$

$$Y_g^* = (F_g' Y_g)$$

$$\epsilon_g^* = (F_g' \epsilon_g)$$

Therefore the likelihood function for the transformed model is:

$$p(D|\beta, \sigma, \lambda) = (2\pi\sigma^2)^{(-n/2)} \sum_{g=1}^{G} \mid A_g^* \mid \times$$

$$exp\left(-\frac{1}{2\sigma^2}\sum_{g=1}^{G}((A_g^* Y_g^* - Z_g^*\beta)'(A_g^* Y_g^* - Z_g^*\beta))\right)$$

The prior distribution could be written as

$$\pi(\beta, \sigma_0^2, \lambda) = \pi(\beta \mid \sigma_0^2)\pi(\sigma_0^2)\pi(\lambda) = N_k(\mu, \sigma_0^2 T) IG(\alpha, \delta) U(1/\rho_{min}, 1/\rho_{max})$$

$$= \frac{1}{(2\pi)^{k|2} \mid T \mid^{1|2} (\sigma_0^2)^{k|2}} \times exp(-\frac{1}{2\sigma_0^2})(\beta - \mu)' T^{-1}(\beta - \mu)$$

$$\times \frac{\delta^\alpha}{\Gamma(\alpha)}(\sigma_0^2)^{-(\alpha+1)} exp\left(\frac{-\delta}{\sigma_0^2}\right) \times \pi(\lambda)$$

$$= \frac{\delta^{\alpha}}{(2\pi)^{k/2}|T|^{1/2}\Gamma(\alpha)}(\sigma_0^2)^{-(\alpha+(k/2)+1)})\times$$

$$\exp\left[-\frac{\left((\beta-\mu)' T^{-1}(\beta-\mu)+2\delta\right)}{2\sigma_0^2}\right]\pi(\lambda)$$

Therefore, the posterior distribution for the model takes the form:

$$p(\beta, \sigma^2, \lambda \mid D) = \frac{p(D \mid \beta, \sigma^2, \lambda)\pi(\beta, \sigma^2)\pi(\lambda)}{p(D)}$$

$$p(\beta, \sigma^2, \lambda \mid Y^*, W^*, Z^*) \propto (\sigma^2)^{-(\alpha^* + (k/2)+1)} \sum_{g=1}^{G} \mid A_g^* \mid \times$$

$$exp\left(-\frac{1}{2\sigma^2}\left[(\beta - \mu)' T^{-1}(\beta - \mu) + 2\delta\right]\right) \times$$

$$exp\left(-\frac{1}{2\sigma^2}\sum_{g=1}^{G}((A_g^* Y_g^* - Z_g^* \beta)'(A_g^* Y_g^* - Z_g^* \beta))\right) \times \pi(\lambda)$$

Then the conditional posteriors distributions are:

$$p(\beta|\lambda, \sigma_0^2) \sim N(\beta^*, \sigma_0^2 T^*)$$

where

$$\beta^* = (\sum_r^R Z^{*'} Z^* + T^{-1})^{-1})(Z^* A^* Y* + T^{-1}\beta)$$

$$T^* = (\sum_{r=1}^R (Z^{*'} Z* + T^{-1})^{-1}$$

and

$$p(\sigma_0^2|\beta,\lambda) \sim IG(\alpha^*,\delta^*)$$

where

$$\alpha^* = a + n/2$$

$$\delta^* = b + (\beta^{'} T^{-1}\beta + \sum_{r=1}^{R} Y^{*'} A^{*'} A^{*} Y - (\beta_n^{'}\Sigma_n^{-1}\beta_n))/2$$

$$p(\lambda|\beta,\sigma) \propto \frac{p(\lambda,\beta,\sigma|D)}{p(\beta,\lambda,\sigma)}$$

$$\propto |I_n - \lambda w| exp\left(-\frac{1}{2\sigma^2}(A^*Y^* - Z^*\beta)'(A^*Y^* - Z^*\beta)\right)$$

Simulation setting

- All groups are assumed to have different sizes (between: 2 and 10, 2 and 15, 2 and 30, and, 2 and 50).
- The number of groups are set to 67 and 102.
- The data generating process are specified as follows : $\lambda = 0.5$, $\beta_{11} = 1$, $\beta_{12} = -1$, $\beta_{21} = 1$, $\beta_{22} = -1$ and $\sigma = 1$.
- The data generation process for each of the X_g*variables* is $N_{m_g}(0, I_{m_g})$
- The continuous dependent variable Y_g can be directly generated based on the model.
- The number of simulations are 100.

Simulation setting

- In particular, we use vague prior distributions setting $\beta_0 = 0, \Sigma = \textit{diag}(1000), \alpha = 0.001$ and $\delta = 0.001$.
- We sampled directly β and σ_0^2 from Gibbs sampling steps since their posterior conditional distributions have closed forms
- Nevertheless, we use the Metropolis–Hastings (M–H) algorithm for sampling the social interaction parameter λ because the full conditional distribution for λ is nonstandard due to the presence of W_r.

Table: Root Square Mean Error (RMSE)

Group size	Num gr.	Approach	Lambda	Beta_11	Beta_12	Beta_21	Beta_22	Sigma
2 to 10	67	ML	0.3344	0.0796	0.0850	0.2546	0.2677	0.0687
		Bayes	0.2405	0.0626	0.0747	0.2647	0.2743	0.0599
	102	ML	0.2565	0.0588	0.0677	0.1771	0.2080	0.0523
		Bayes	0.1911	0.0519	0.0534	0.2199	0.1981	0.0510
2 to 15	67	ML	0.3138	0.0559	0.0590	0.2975	0.3413	0.0480
		Bayes	0.2156	0.0554	0.0508	0.2752	0.2871	0.0395
	102	ML	0.2297	0.0476	0.0480	0.2526	0.2463	0.0375
		Bayes	0.1906	0.0466	0.0434	0.2166	0.2388	0.0321
2 to 30	67	ML	0.3435	0.0416	0.0411	0.3438	0.3469	0.0346
		Bayes	0.2226	0.0326	0.0330	0.3678	0.3506	0.0267
	102	ML	0.2961	0.0287	0.0286	0.2871	0.2949	0.0240
		Bayes	0.2030	0.0311	0.0266	0.2725	0.2897	0.0240
2 to 50	67	ML	0.4567	0.0326	0.0257	0.4665	0.4176	0.0214
		Bayes	0.2276	0.0264	0.0276	0.4346	0.4604	0.0188
	102	ML	0.3319	0.0251	0.0239	0.3666	0.3374	0.0175
		Bayes	0.1878	0.0195	0.0218	0.3580	0.3201	0.0151

Table: Mean Absolute Error (MAE)

Group size	Num gr.	Approach	Lambda	Beta_11	Beta_12	Beta_21	Beta_22	Sigma
2 to 10	67	ML	0.2395	0.0617	0.0660	0.1963	0.2131	0.0534
		Bayes	0.2071	0.0516	0.0600	0.2028	0.2240	0.0488
	102	ML	0.1990	0.0490	0.0556	0.1400	0.1626	0.0425
		Bayes	0.1618	0.0437	0.0424	0.1723	0.1518	0.0397
2 to 15	67	ML	0.2398	0.0439	0.0485	0.2394	0.2641	0.0389
		Bayes	0.1904	0.0453	0.0395	0.1993	0.2348	0.0333
	102	ML	0.1850	0.0364	0.0386	0.1845	0.1907	0.0307
		Bayes	0.1614	0.0372	0.0351	0.1852	0.1936	0.0266
2 to 30	67	ML	0.2638	0.0341	0.0329	0.2614	0.2770	0.0258
		Bayes	0.1833	0.0264	0.0264	0.2906	0.2743	0.0212
	102	ML	0.2179	0.0234	0.0231	0.2316	0.2282	0.0187
		Bayes	0.1697	0.0252	0.0221	0.2134	0.2440	0.0190
2 to 50	67	ML	0.3463	0.0269	0.0200	0.3604	0.3380	0.0167
		Bayes	0.1832	0.0217	0.0225	0.3334	0.3662	0.0150
	102	ML	0.2454	0.0208	0.0189	0.2853	0.2685	0.0140
		Bayes	0.1538	0.0155	0.0175	0.2884	0.2486	0.0122

Introduction

- One interesting application for this type of models is bank profitability.
- The role of banks remains central in various important aspects for the economy
- Most of studies on bank profitability, estimate the impact of numerous factors that may be important in explaining profits using linear models
- Even though these studies show that it is possible to conduct a meaningful analysis of bank profitability, there are still some issues to be addressed.

Model Specification

- Therefore, we propose the following model specification:

$$\begin{aligned} Profit_g = \lambda W_g Profit_g + CashF_g \beta_1 + WorKC_g \beta_2 + Leverage_g \beta_3 \\ + EBITDAint \beta_4 + ROA \beta_5 + W_g CashF_g \beta_6 + W_g WorkC_g \beta_7 \\ + W_g Leverage_g \beta_8 + W_g EBITDAint_g \beta_9 + W_g ROA_g \beta_{10} + I_{m_g} \alpha_g + \epsilon_g \end{aligned}$$

Table: Bayesian Estimation

Variables	Bayesian M-H Mean	2.50%	25%	summary 50%	75%	97.50%
Lambda	0.9986	0.9946	0.9980	0.9990	0.9996	1.0000
Sigma	1.053	1.0440	1.0500	1.0530	1.0560	1.0620
WK	-0.0257	-0.0486	-0.0339	-0.0256	-0.0178	-0.0025
Cash_Flow	-0.0071	-0.0270	-0.0148	-0.0073	-0.0003	0.0152
Leverage	0.0296	0.0093	0.0223	0.0294	0.0370	0.0487
EBITDA_int	0.0335	0.0150	0.0276	0.0337	0.0397	0.0512
ROA	-0.1039	-0.1243	-0.1110	-0.1038	-0.0970	-0.0840
Lag WK	-0.0364	-0.1718	-0.0783	-0.0404	0.0086	0.0838
Lag Cash Flow	-0.3436	-0.1348	-0.0825	-0.0372	0.0022	0.1046
Lag Leverag	-0.0111	-0.1134	-0.0447	-0.0069	0.0299	0.0714
Lag EBITDA_int	0.0512	-0.0687	-0.0060	0.0519	0.0993	0.1831
Lag ROA	-0.1142	-0.2379	-0.1470	-0.1104	-0.0752	-0.0125

We welcome your comments, questions, and suggestions!!!