A New Latent Class to Fit Spatial Econometrics Models with Integrated Nested Laplace Approximations

Virgilio Gómez-Rubio,[1] Roger Bivand[2,3] and Håvard Rue[4]

1. University of Castilla-La Mancha, 02071 Albacete, Spain
2. Norwegian School of Economics, 5045 Bergen, Norway
3. Adam Mickiewicz University, 61-680 Poznań, Poland
4. Norwegian University for Science and Technology, 7491 Trondheim, Norway

NTNU – Trondheim
Norwegian University of Science and Technology

The new *slm* latent model for estimating spatial econometrics models using INLA (Rue et al., 2009) has recently been introduced. It will be described briefly and its use will be demonstrated for Gaussian and non-Gaussian dependent variables.

Spatial Econometrics Models

- For a review of spatial econometrics models, see Anselin (2010) and LeSage (2014).
- We have a vector y of observations from n different regions; the adjacency structure of these regions is available in a matrix W (here row-standardised).
- All these models can be rewritten so that the response y only appears on the left hand side. For example, the spatial error model (SEM) is:

$$y = X\beta + (I_n - \rho_{\mathrm{Err}} W)^{-1} e; \, e \sim MVN(0, \sigma^2 I_n); \quad (1)$$

and the spatial lag model (SLM):

$$y = (I_n - \rho_{\mathrm{Lag}} W)^{-1}(X\beta + e); \; e \sim MVN(0, \sigma^2 I_n); \quad (2)$$

- Bivand et al. (2014, 2015) described how INLA could be used to fit some spatial econometrics models, conditioning on values of the spatial autocorrelation parameter, and later combining using Bayesian model averaging to obtain the posterior marginals of the parameters of the desired model.

INLA: the slm model

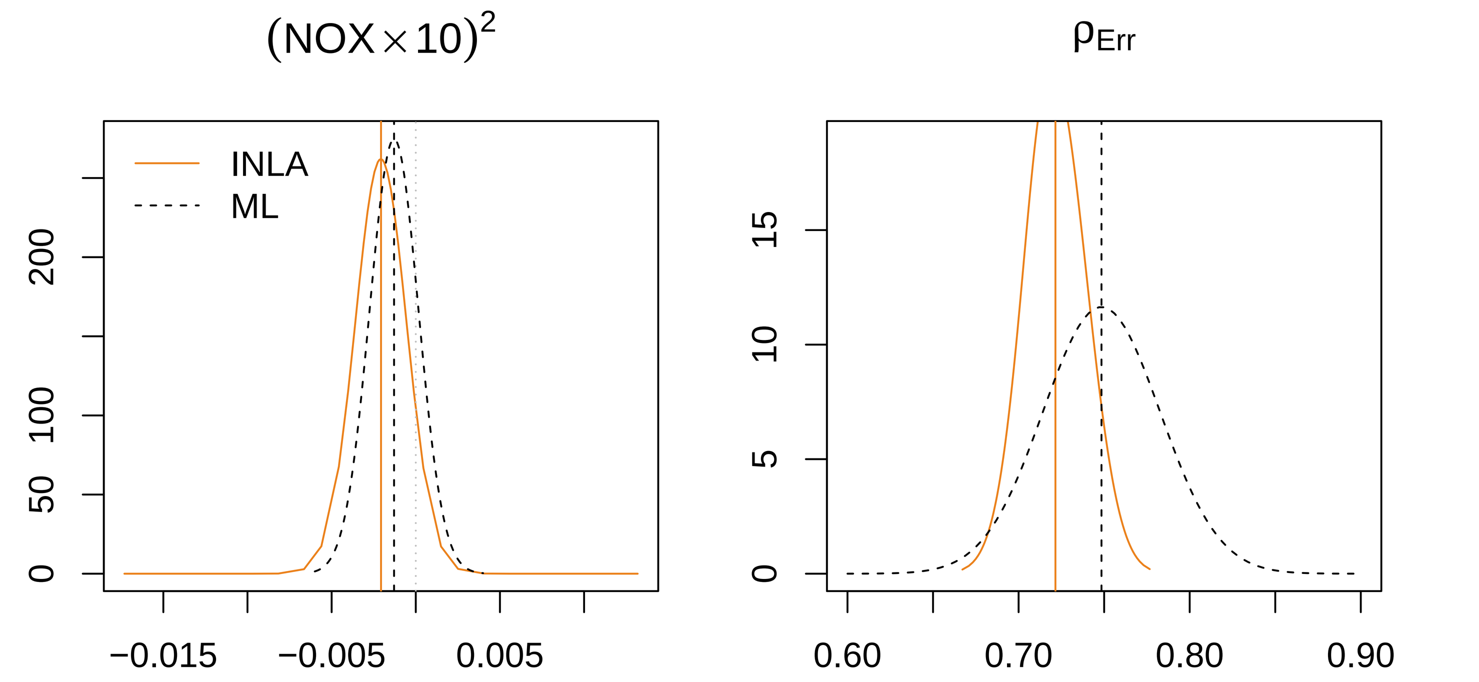
- The new `slm` latent model makes the use of Bayesian model averaging unnecessary for single spatial parameter models, and implements the following expression as a random effect that can be included in the linear predictor:

$$\mathbf{x} = (I_n - \rho W)^{-1}(X\beta + \varepsilon) \quad (3)$$

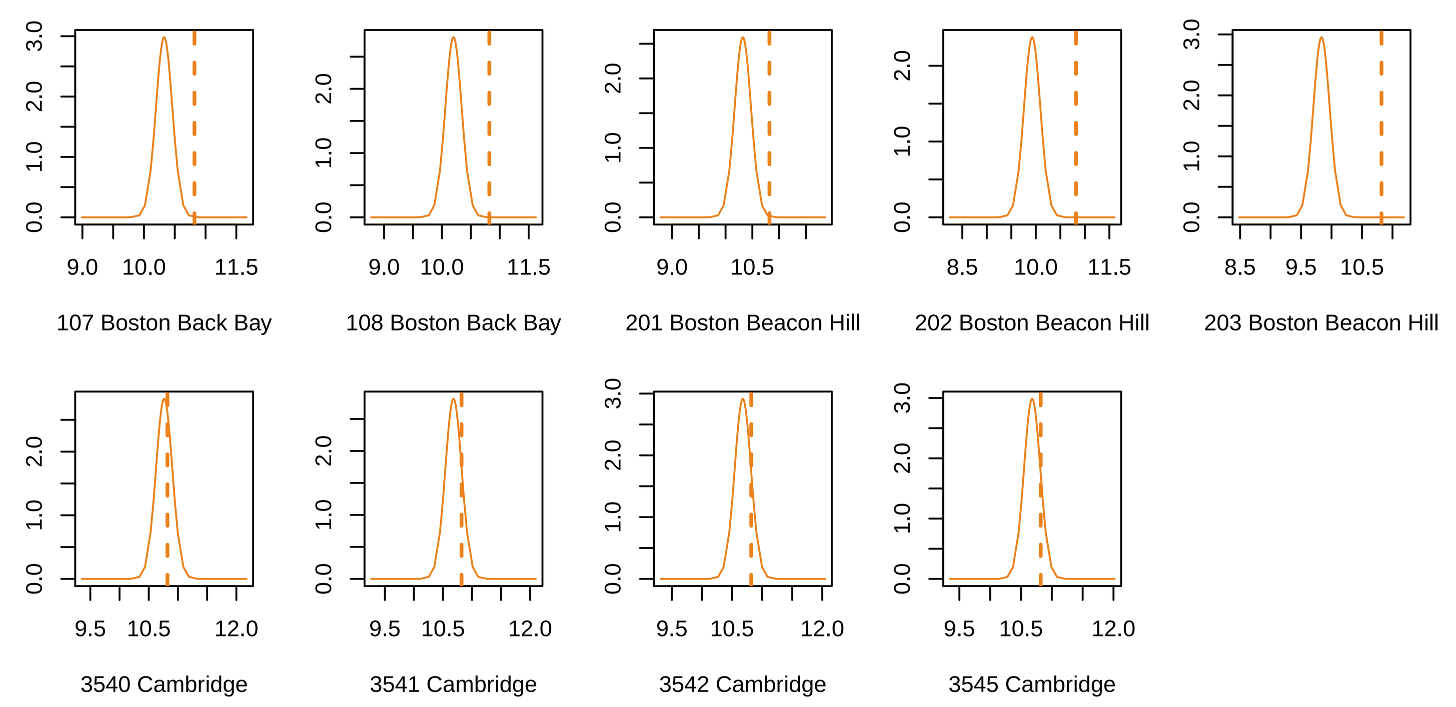
- Here, $\mathbf{x}$ is a vector of n random effects, I_n is the identity matrix of dimension $n \times n$, ρ is a spatial autocorrelation parameter, W is a $n \times n$ weight matrix, X a matrix of covariates with coefficients β and ε is a vector of independent Gaussian errors with zero mean and precision τI_n.
- In this latent model, we need to assign prior distributions to the vector of coefficients β, spatial autocorrelation parameter ρ and precision of the error term τ.
- By default, β takes a multivariate Gaussian distribution with zero mean and precision matrix Q (which must be specified); $\mathrm{logit}(\rho)$ takes a Gaussian prior with zero mean and precision 10; and $\log(\tau)$ takes a log-gamma prior with parameters 1 and $5 \cdot 10^{-5}$.
- Note that many spatial econometrics models can be derived from this implementation; in particular, the SEM model is a particular case with $\beta = 0$; the SLM model can be fitted with no modification.
- Details of the implementation are available at: `http://www.math.ntnu.no/inla/r-inla.org/doc/latent/slm.pdf`.

Comparative results: Boston

- First, we use the study by Harrison and Rubinfeld (1978) of the median value of owner-occupied houses in the Boston area using 13 covariates; the median value has been censored at $50,000 and we omit 16 tracts that are censored, leaving 490 observations (see Pace and Gilley, 1997).



- Figure 1 shows coefficient estimates and marginals for the air pollution covariate: $(\mathrm{NOX} \times 10)^2$, and the spatial error coefficient ρ_{Err}. The ML point estimates and standard errors agree well with the INLA SEM marginals and means.



- Figure 2 shows INLA SEM marginals of fitted values for 9 censored tracts; the censoring boundary shown by dashed vertical line. The censored tracts in Cambridge neighbour similar tracts, but in Boston, the covariates and neighbouring tracts do not imply high house values (here in logs of 1970 USD).

Comparative results: New Orleans

- LeSage et al. (2011) study the probabilities of reopening a business in New Orleans in the aftermath of Hurricane Katrina; here we reproduce the analysis with a continuous probit link function.

	mean	0.025quant	0.5quant	0.975quant
(Intercept)	-19.8032	-31.7854	-19.7124	-8.3258
flood depth	-0.4322	-0.6046	-0.4295	-0.2746
log medinc	1.9251	0.8069	1.9165	3.0929
small size	-0.3475	-0.7464	-0.3471	0.0485
large size	-0.4084	-1.3410	-0.4039	0.4990
low status customers	-0.3326	-0.8481	-0.3342	0.1916
high status customers	0.1372	-0.2378	0.1371	0.5128
owntype sole proprietor	0.7766	0.2358	0.7744	1.3290
owntype national chain	0.0974	-0.9340	0.1019	1.1026
Rho for idx	0.5982	0.3782	0.6102	0.7518

- Table 1 shows the results of fitting a spatial error probit model; the middle levels of categorical covariates business size, customer status, and ownership type are included in the intercept.

Work in progress includes the impacts/effects marginals when the spatially lagged response is included in the model, and the use of Bayesian model averaging when more than one spatial coefficient is included.

To download code and references, please visit:
`http://spatial.nhh.no/misc/ss15/INLA_slm`