Bayesian spatial econometrics: a software architecture

Nikolas Kuschnig (`nkuschnig@wu.ac.at`)

June 23, 2022

Department of Economics
Vienna University for Economics and Business

Motivation

Bayesian approaches can play a crucial role for the **development of spatial econometric methods**. For example

- Bayesian model averaging [7, 16],
- hierarchical models [6, 9, 14],
- flexible treatment of connectivity [8, 11],
- limited dependent variables [12, 15],

However, they are **rarely used in applied work**.

Bayesian spatial software

Implementations of Bayesian spatial methods need to accommodate

- the **flexibility** of Bayesian modelling,
- the **peculiarities** of spatial econometrics.

Domain-specific packages, such as **spatialreg** [3], are constrained in their flexibility, while general purpose software like **Stan** [4] can be inefficient and ineffective [21].

Overview

In this talk, I present a **software architecture** for Bayesian spatial econometric methods that is

- **powerful and extensible**,
- while remaining **accessible** and
- **easy to maintain**.

I will demonstrate an *implementation* in the **bsreg** package [13], and may briefly present some applied results that highlight a need for such software.

A brief recap of spatial econometrics and Bayesian estimation

Spatial econometric models

A **general** spatial econometric model is given by

$$\mathbf{y} = \boxed{\lambda_1 \mathbf{W}\mathbf{y}} + \mathbf{X}\boldsymbol{\beta} + \boxed{\mathbf{W}\mathbf{X}\boldsymbol{\theta}} + \mathbf{u},$$
$$\mathbf{u} = \boxed{\lambda_2 \mathbf{W}\mathbf{u}} + \boldsymbol{\varepsilon},$$

where $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$. We use the $n \times n$ matrix $\mathbf{W}$ to induces spatial lags that extend the standard linear model. The lags allow for

1. spatial autoregressive behaviour,
2. spatial interference from explanatories,
3. spatially autocorrelated errors.

Bayesian estimation

We can estimate spatial models with MCMC **sampling methods**. These methods work by isolating conditional posteriors, such as $p(\lambda_1 | \mathbf{y}, \beta, \sigma^2)$, and drawing from them. It helps to re-interpet the models, e.g. as

$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

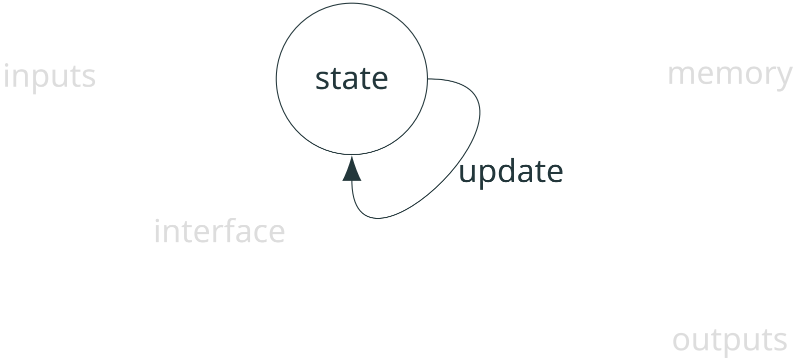
$$\mathbf{y} = \boxed{\mathbf{S}(\lambda | \mathbf{W})}^{-1} \mathbf{z},$$

- with a **latent** variable $\mathbf{z}$ (and a nested linear model),
- a non-linear **spatial filter** $\mathbf{S} = (\mathbf{I} - \lambda \mathbf{W})$.

Bayesian spatial econometrics — a software architecture

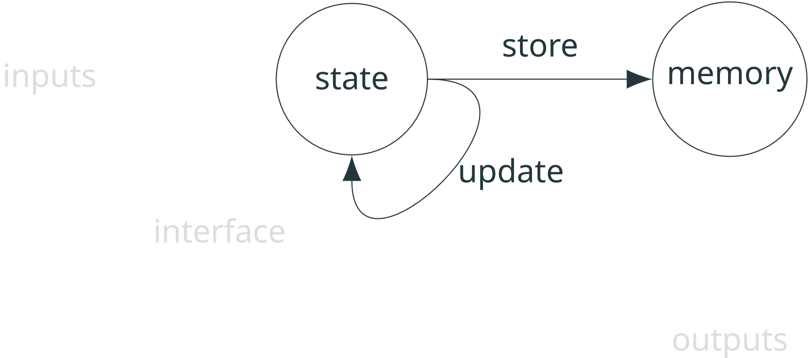
In the beginning was a prototype I

We start by oberving that our MCMC samplers are fully characterised by a *state*, and *rules to update* the state — i.e. **state machines**.



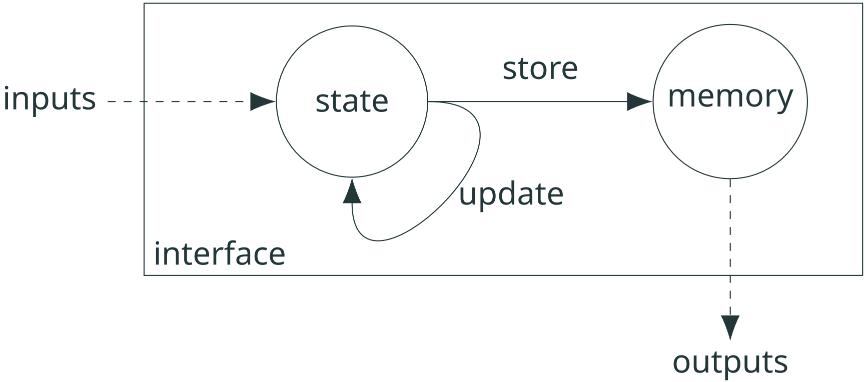
In the beginning was a prototype II

For the state machine to be useful, we want to update and **store its state** — the posterior draws — repeatedly.



In the beginning was a prototype III

Finally, we want to **interact with the state machine** (initialise and adjust it), and use the outputs via an **interface**.

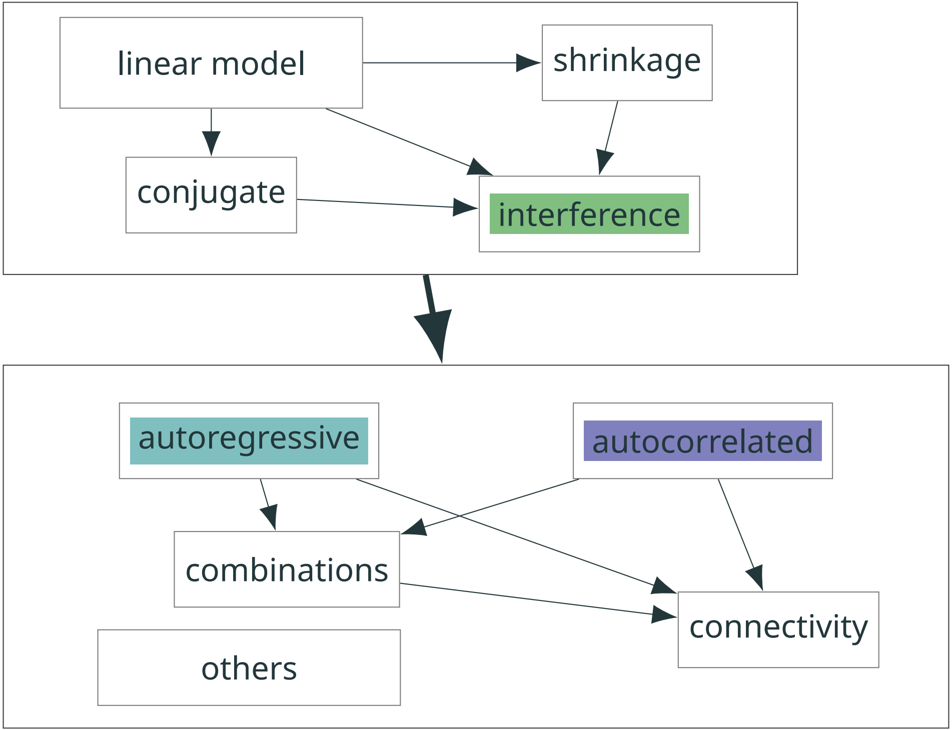


An object-oriented system

Our core is a prototype-based **object-oriented** (OO) system around this state machine.

We depart from a simple linear model that we can readily adjust (priors, computation) and extend — e.g. with

- local spatial lags,
- spatial filters, and
- parameterised connectivity.


linear model
shrinkage
conjugate
interference
autoregressive
autocorrelated
combinations
connectivity
others

Functional object-orientation

A **functional** approach within the object-oriented system affords us *flexibility and extensibility*. Only the overall structure is predetermined.

To make this setup practical, we use a **programming interface** to

- instantiate objects (choose priors, prepare composite objects),
- work with objects (tune samplers, obtain samples),
- prepare outputs.

Hybrid programming for hybrid models

Figure 2: La Chimera di Arezzo (CC BY-SA 3.0 by Wikimedia User Sailko).

A user interface

Finally, we use an **idiomatic user interface**, connecting the *programming interface* and the *language*.

For R this might include

- the formula interface — e.g. `bm(attention ~ colors)`,
- a `plot()` method for visualisation,
- a `summary()` method to *summarise*,
- interfaces to **coda** [17], **broom** [19], **ggplot2** [20], etc.

Implementation in R

The bsreg package

In the **bsreg** package [13] I try to implement this architecture in R.

I use the **R6** object-oriented system [5] — which is a more classical OO system than e.g. **S4** [18] — to represent

- a **range of models**, roughly as
  - skeleton > non-conjugate linear model >
  - conjugate prior | shrinkage priors >
  - local spatial lags > filters > connectivity,
- Metropolis-Hastings samplers.

A functional design for intermediate steps ensures flexiblity.

Using bsreg

To make **bsreg** accessible, I provide two interfaces.

1. A programming interface
    - `o <- set_options(set_SLX(delta_scale = .1))`
    - `m <- get_bslx(y = y, X = X, options = o)`
    - `s <- sample(m, n_burn = 1000) # modification in place`

Using bsreg

To make **bsreg** accessible, I provide two interfaces.

1. A programming interface
   - `o <- set_options(set_SLX(delta_scale = .1))`
   - `m <- get_bslx(y = y, X = X, options = o)`
   - `s <- sample(m, n_burn = 1000) # modification in place`
2. A user interface
   - `m <- blm(log(sales) ~ log(price), data = d)`
   - `summary(m) # omitted spillovers?`
   - `m <- bslx(log(sales) ~ log(price), W = W_id, data = d)`
   - `plot(m) # not converged?`
   - `m <- bm(m, n_save = 5000)`

An applied demonstration

Cigarette demand elasticity

Consider the following panel model of US cigarette demand [1]

$$\ln C_{it} = \alpha + \beta_{price} \ln P_{it} + \beta_{income} \ln I_{it} + \mu_i + \phi_t + \varepsilon_{it}.$$

There is good evidence for **spillover effects between states**, so a spatial econometric model may be a sensible choice. Let's investigate the effect of income.

Average partial effects

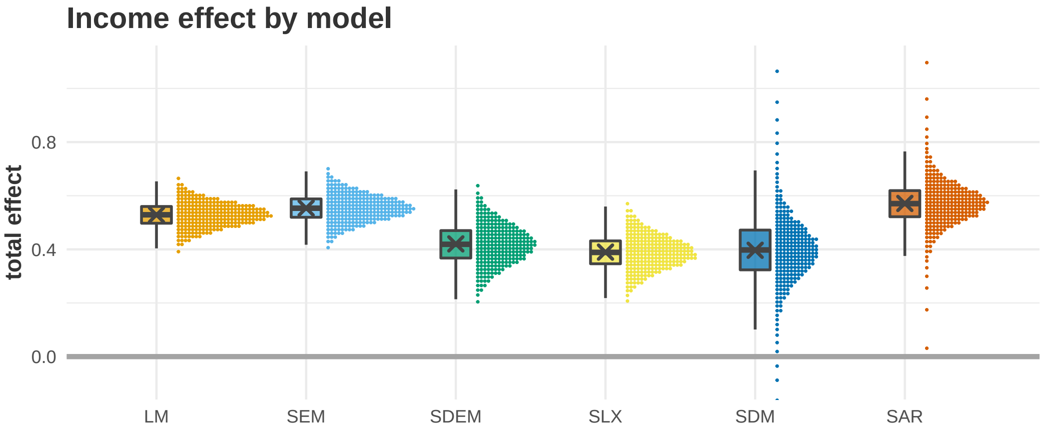


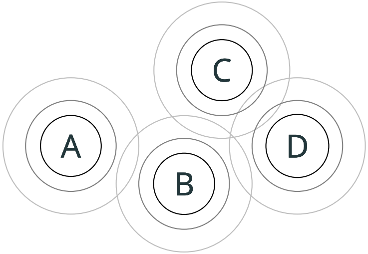
Figure 3: Average income effect using a linear model and spatial models with contiguity-based $\mathbf{W}$.

Uncertainty in spatial econometrics

There is **uncertainty around spillover effects**, but also connectivity.

The preferred model of Halleck-Vega and Elhorst [10] is a SLX model (with a local spatial lag) with **parameterised connectivity**. They use an inverse-distance decay matrix, where

- $[\mathbf{W}]_{ij} = d_{ij}^{-\delta}$ for $i \neq j$ and 0 otherwise,
- with distance d and free parameter δ.



Uncertainty around connectivity

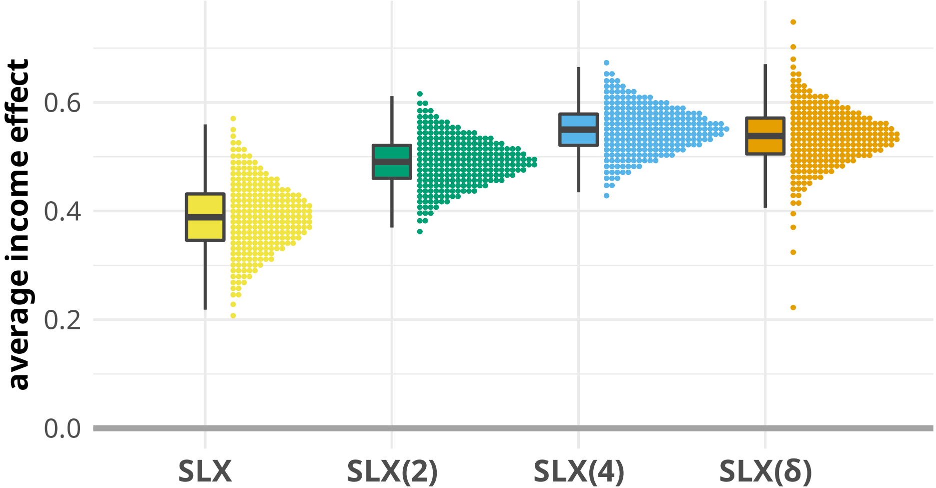


Figure 4: Average income effect using contiguity-based $\mathbf{W}$, and inverse-distance decay-based $\mathbf{W}$ (scaled with the maximum eigenvalue) with fixed $\delta \in \{2, 4\}$ and parameterised δ.

Two final scaling issues

First, posterior draws reflect **uncertainty around both** δ and θ at the same time, while the Delta method [10] holds δ fixed.

Two final scaling issues

First, posterior draws reflect **uncertainty around both** δ and θ at the same time, while the Delta method [10] holds δ fixed.

Second, consider the average partial effects of the SLX model

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}_k} = \mathbf{I}\beta_k + \mathbf{W}\theta_k$$

When $\mathbf{W}$ is row-stochastic (or generally $\sum_{i=1}^{n}\sum_{j=1}^{n}[\mathbf{W}]_{ij} = n$) the parameter θ gives us the *average partial effect*.

Two final scaling issues

First, posterior draws reflect **uncertainty around both** δ and θ at the same time, while the Delta method [10] holds δ fixed.

Second, consider the average partial effects of the SLX model

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}_k} = \mathbf{I}\beta_k + \mathbf{W}\theta_k$$

When $\mathbf{W}$ is row-stochastic (or generally $\sum_{i=1}^{n} \sum_{j=1}^{n} [\mathbf{W}]_{ij} = n$) the parameter θ gives us the *average partial effect*.

This is **not the case in general** — Halleck-Vega and Elhorst [10] and the **spatialreg** package [2] *do not account for this* and **overestimate** spillover effects **by** a factor of **seven**.

Thank you!

Find me, the paper, and the package at:

`<kuschnig.eu>`

`<doi.org/10.1007/s43071-022-00023-w>`

`<CRAN.R-project.org/pkg=bsreg>`

References i

[1] Badi H Baltagi and Dong Li.
Prediction in the panel data model with spatial correlation.
In *Advances in Spatial Econometrics*, pages 283–295. Springer, 2004.

[2] Roger Bivand, Jan Hauke, and Tomasz Kossowski.
Computing the Jacobian in Gaussian spatial autoregressive models: an illustrated comparison of available methods.
Geographical Analysis, 45(2):150–179, 2013.

[3] Roger S. Bivand, Edzer Pebesma, and Virgilio Gómez-Rubio.
***Applied spatial data analysis with R*.**
Springer, NY, 2013.

[4] Bob Carpenter, Andrew Gelman, Matthew Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell.
Stan: a probabilistic programming language.
Journal of Statistical Software, 76(1):1–32, 2017.

References ii

[5] Winston Chang.

***R6: encapsulated classes with reference semantics*, 2021.**

R package version 2.5.1.

[6] Gary J. Cornwall and Olivier Parent.

Embracing heterogeneity: the spatial autoregressive mixture model.

Regional Science and Urban Economics, 64:148–161, 2017.

[7] Jesús Crespo Cuaresma and Martin Feldkircher.

Spatial filtering, model uncertainty and the speed of income convergence in Europe.

Journal of Applied Econometrics, 28(4):720–741, 2013.

[8] Nicolas Debarsy and James P. LeSage.

Bayesian model averaging for spatial autoregressive models based on convex combinations of different types of connectivity matrices.

Journal of Business & Economic Statistics, pages 1–12, 2020.

[9] Guanpeng Dong and Richard Harris.
Spatial autoregressive models for geographically hierarchical data structures.
Geographical Analysis, 47(2):173–191, 2015.

[10] Solmaria Halleck Vega and J. Paul Elhorst.
The SLX model.
Journal of Regional Science, 55(3):339–363, 2015.

[11] Xiaoyi Han and Lung-Fei Lee.
Bayesian analysis of spatial panel autoregressive models with time-varying endogenous spatial weight matrices, common factors, and random coefficients.
Journal of Business & Economic Statistics, 34(4):642–660, 2016.

[12] Tamás Krisztin, Philipp Piribauer, and Michael Wögerer.
A spatial multinomial logit model for analysing urban expansion.
Spatial Economic Analysis, pages 1–22, 2021.

References iv

[13] Nikolas Kuschnig.
***bsreg: Bayesian spatial regression models*, 2022.**
R package version 0.0.2.

[14] Donald J. Lacombe and Stuart G. McIntyre.
Local and global spatial effects in hierarchical models.
Applied Economics Letters, 23(16):1168–1172, 2016.

[15] James P LeSage.
Bayesian estimation of limited dependent variable spatial autoregressive models.
Geographical Analysis, 32(1):19–35, 2000.

[16] James P. LeSage and R. Kelley Pace.
A matrix exponential spatial specification.
Journal of Econometrics, 140(1):190–214, 2007.

References v

[17] Martyn Plummer.
JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling.
In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, 2003.

[18] R Core Team.
***R: a language and environment for statistical computing*.**
R Foundation for Statistical Computing, Vienna, Austria, 2021.

[19] David Robinson et al.
***broom: convert statistical objects into tidy tibbles*, 2022.**
R package version 0.8.0.

[20] Hadley Wickham.
***ggplot2: elegant graphics for data analysis*.**
Springer-Verlag New York, 2016.

References vi

[21] Levi John Wolf, Luc Anselin, and Daniel Arribas-Bel.

Stochastic efficiency of Bayesian Markov chain Monte Carlo in spatial econometric models: an empirical comparison of exact sampling methods.

Geographical Analysis, 50(1):97–119, 2018.