Bayesian Spatial Econometrics
A Software Architecture

Nikolas Kuschnig (nkuschnig@wu.ac.at)
Vienna University of Economics and Business

I. Introduction

Bayesian approaches play an important role in the development of spatial econometrics, but are **uncommon in applied work**. This is, at least in part, due to a lack of accessible and flexible software. General-purpose software struggles with spatial particularities [6], while classical implementations do not harness the flexibility of Bayesian modelling.

I present a **layered, objected-oriented software architecture** [5, 4], which combines accessibility with extensibility. The interface is split into a programming and user interface, facilitating flexible and maintainable code that is easy-to-use. I implement this approach in the **bsreg** R package, and demonstrate the flexibility of the Bayesian approach using a well-known dataset on cigarette demand.

II. Methods

Spatial econometrics extends the linear regression model

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

where $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{X} \in \mathbb{R}^{n \times k}$, and $\mathbf{e} \in \mathbb{R}^n$ is an error term with mean zero. We include **spatial lags** via the **known connectivity matrix** $\mathbf{W} \in \mathbb{R}^{n \times n}$ (with elements $\mathbf{w}_{ij} > 0$ for neighbours i and j, where $i \neq j$, and 0 otherwise). A general spatial econometric model is given by

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

In practice, one or two lags are used; they convey

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

The interference is *linear*; autoregressive and autocorrelated terms can be understood in terms of a **latent variable z** and a **spatial filter S** $= (\mathbf{I} - \lambda\mathbf{W})$, e.g.

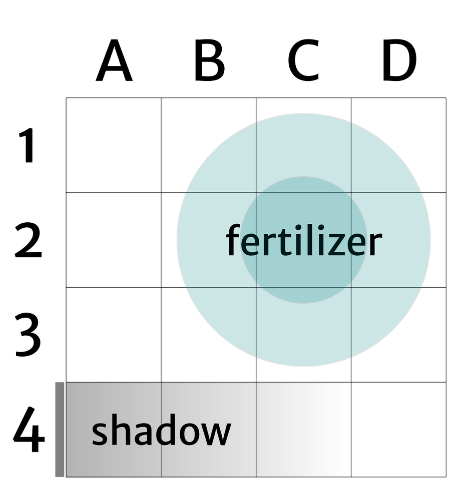
$$\mathbf{z} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$
$$\mathbf{y} = \mathbf{S}(\lambda|\mathbf{W})^{-1}\mathbf{z}.$$

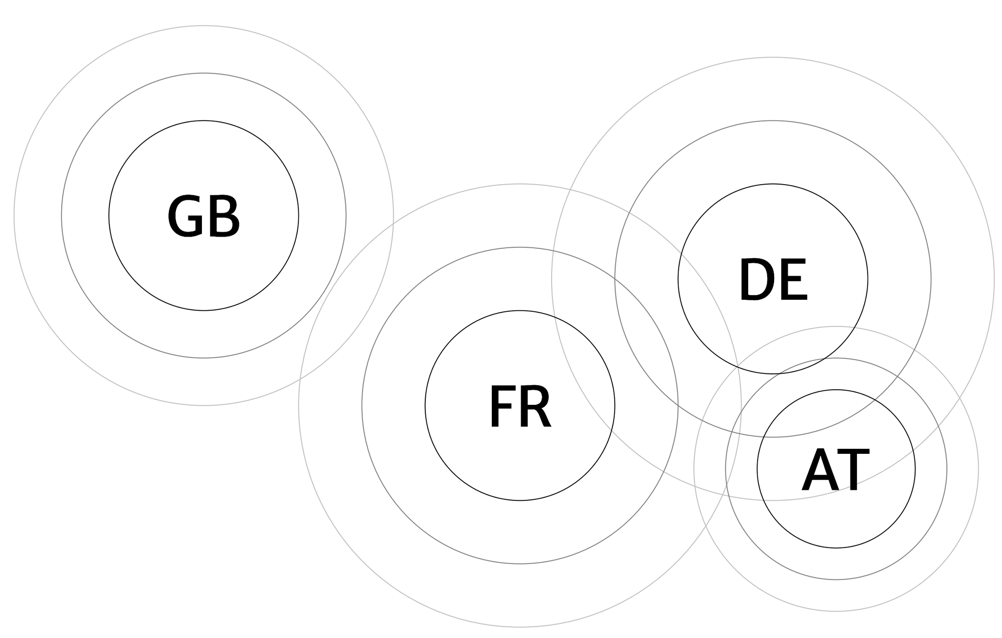
III. Specificities

Two domain-specific aspects are particularly important for useful software implementations.

1. We need to effectively deal with **spatial information**, ideally incorporating it further into the model.
2. With a spatial filter, we need to evaluate the **Jacobian** $|\mathbf{S}|$, which is computationally expensive [1], hindering generalist software [6].

This motivates a domain-specific package that can be embedded in the existing R infrastructure.





IV. Architecture

The programming interface uses object-oriented prototypes; samplers are represented as simple **state machines**. This facilitates extensions to shrinkage setups (e.g. the Horseshoe), but also spatial models (with modelled connectivity), limited dependent variable models, or stochastic volatility.

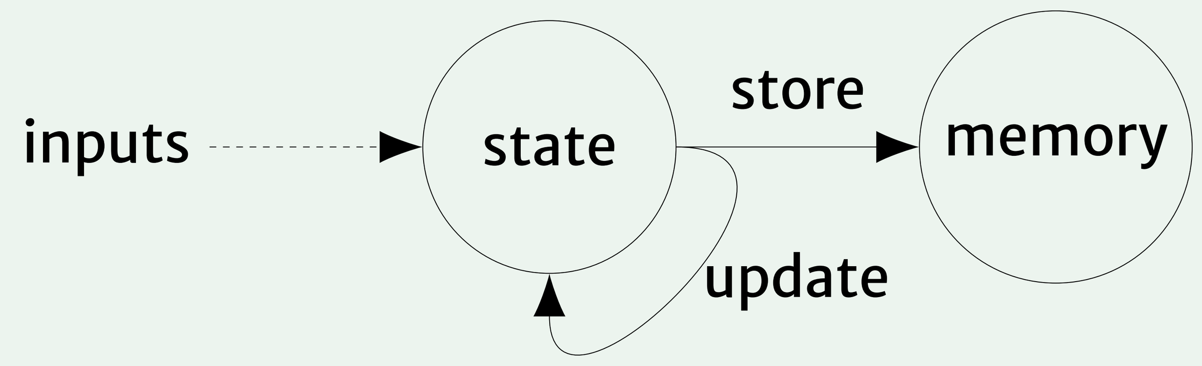


Figure: An idealised object with Markov chain Monte Carlo sampler and storage.

The **bsreg** package implements this structure using **R6** [2], coupled with the familiar formula interface:

```
bslx(log(sales) ~ log(price), data = cigarettes, W = W)
```

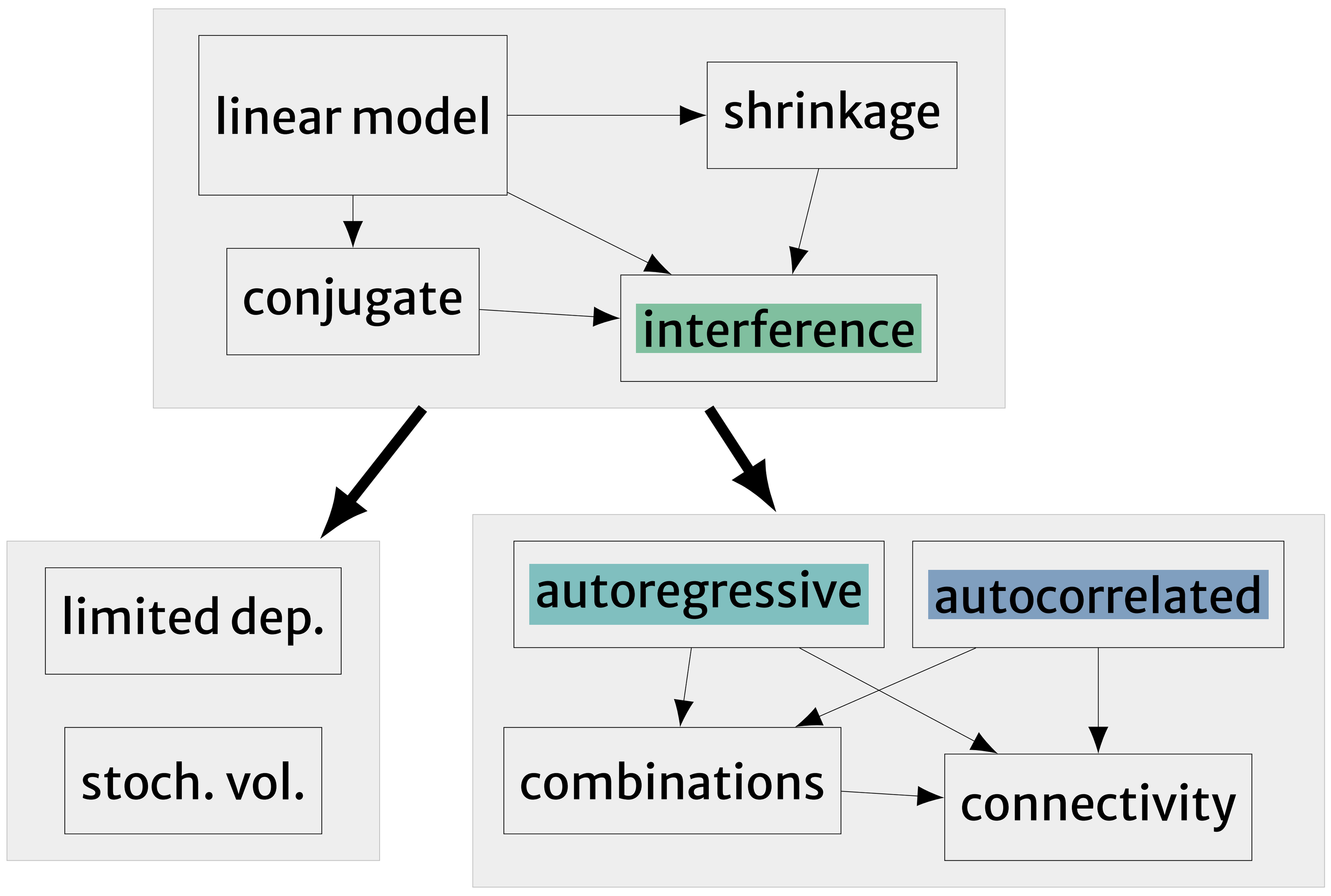


Figure: The inheritance structure of the object-oriented architecture.

V. Demonstration

To demonstrate, I revisit an analysis of cross-state spillover effects in US cigarette demand [3]. I find that

→ modelling connectivity (via a distance-decay parameter) and using posterior intervals **better reflects uncertainty**,

→ **spillovers were previously misjudged**, due to neglect of the *scale* implied by **W** (in scripts and software).

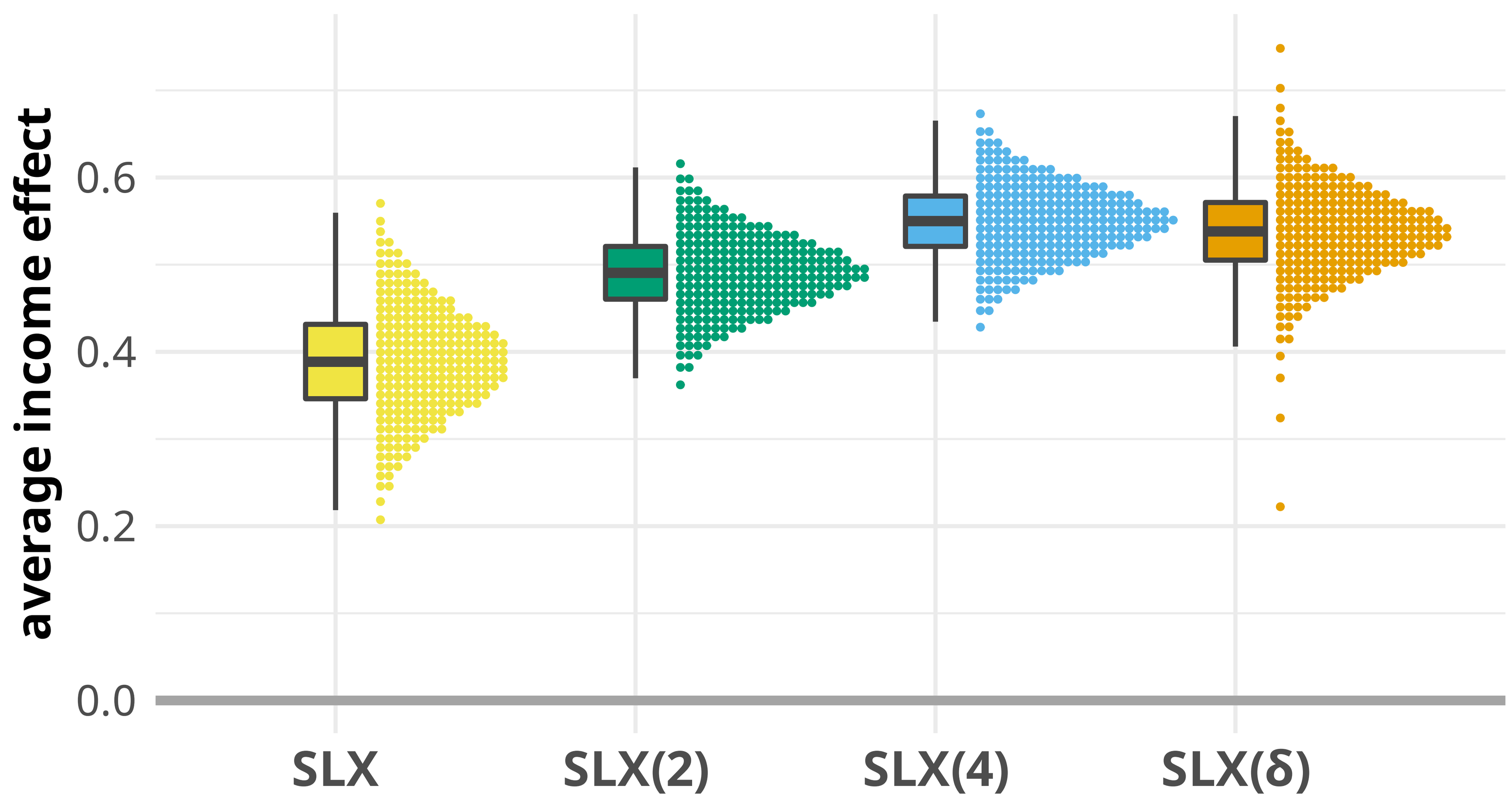


Figure: Posterior impacts on cigarette demand with contiguity- and distance-based **W**.

Contact

Find my work online at kuschnig.eu (QR code); contact me via mail at nkuschnig@wu.ac.at or messenger pigeon at twitter.com/__nkuschnig.

[1] Bivand, R. et al, 2013. Computing the Jacobian in Gaussian spatial autoregressive models: an illustrated comparison of available methods. *GA*, 45 (2). doi:10.1111/gean.12008.

[2] Chang, W. ***R6**: encapsulated classes with reference semantics*, 2021. URL CRAN.R-project.org/package=R6.

[3] Halleck Vega, S. et al, 2015. The SLX model. *JRS*, 55 (3). doi:10.1111/jors.12188.

[4] Kuschnig, N., 2022. Bayesian spatial econometrics: a software architecture. *JSE*, (forthcoming). doi:10.1007/s43071-022-00023-w.

[5] Kuschnig, N. ***bsreg**: Bayesian spatial regression models*, 2022. URL CRAN.R-project.org/package=bsreg.

[6] Wolf, L.J. et al, 2018. Stochastic efficiency of Bayesian Markov chain Monte Carlo in spatial econometric models: an empirical comparison of exact sampling methods. *GA*, 50(1). doi:10.1111/gean.12135.