

# Understanding the LCAR Model

## Intuition, notation, and the role of lambda

María Barrera

## Motivation

Spatial random effects are often introduced in disease mapping models to account for residual spatial structure that cannot be explained by observed covariates.

Among Conditional Autoregressive (CAR) models, the Leroux CAR (LCAR) formulation is especially appealing because it defines a **continuous spectrum** between independence and full spatial dependence.

The goal of these notes is not to be exhaustive, but to make **explicit** how the parameter  $\lambda$  operates and how the model behaves in relevant limiting cases.

## Joint formulation

Let  $\varphi = (\varphi_1, \dots, \varphi_I)$  denote the spatial random effects. Under the LCAR model:

$$\varphi | \sigma, \lambda \sim \mathcal{N}_I \left( \mathbf{0}_I, \sigma^2 [(1 - \lambda)\mathbf{I}_I + \lambda(\mathbf{D} - \mathbf{W})]^{-1} \right),$$

where:

- $\mathbf{W}$  is the adjacency matrix,
- $\mathbf{D}$  is a diagonal matrix with the number of neighbours for each area,
- $\lambda \in [0, 1]$  controls the strength of spatial dependence.

## Conditional formulation

The full conditional distribution of  $\varphi_i$  is given by:

$$\varphi_i | \varphi_{-i}, \sigma, \lambda \sim \mathcal{N} \left( \frac{\lambda}{1 - \lambda + \lambda m_i} \sum_{j \sim i} \varphi_j, \frac{\sigma^2}{1 - \lambda + \lambda m_i} \right), \quad i = 1, \dots, I.$$

where  $m_i$  denotes the number of neighbours of area  $i$ .

This expression makes explicit how information is shared locally. The conditional mean is a weighted average of neighbouring effects, while the conditional variance adapts to both  $\lambda$  and  $m_i$ .

## Limiting cases

$\lambda = 0$

$$\varphi_i | \varphi_{-i}, \sigma, \lambda = 0 \sim \mathcal{N}(0, \sigma^2),$$

which corresponds to independence between spatial units and homoscedasticity.

$\lambda = 1$

The model reduces to the intrinsic CAR (ICAR) specification, up to the usual sum-to-zero constraint:

$$\varphi_i | \boldsymbol{\varphi}_{-i}, \sigma, \lambda = 1 \sim \mathcal{N} \left( \frac{1}{m_i} \sum_{j \sim i} \varphi_j, \frac{\sigma^2}{m_i} \right).$$

In this case, the conditional mean corresponds to the average of neighbouring effects (maximum smoothing), and the conditional variance decreases as the number of neighbours increases.

$\lambda = 0.5$

$$\varphi_i | \boldsymbol{\varphi}_{-i}, \sigma, \lambda = 0.5 \sim \mathcal{N} \left( \frac{1}{1 + m_i} \sum_{j \sim i} \varphi_j, \frac{2\sigma^2}{1 + m_i} \right).$$

Here, the conditional mean is shrunk towards the neighbourhood average, but does not fully coincide with it. The conditional variance still decreases with  $m_i$ , while allowing for greater local variability compared to the ICAR case.

## Interpretation of $\lambda$

Rather than being fixed by the analyst,  $\lambda$  is estimated from the data and acts as a **regulator of spatial dependence**:

- When data are scarce or noisy, spatial borrowing stabilizes estimates.
- When information is sufficient, the model relies more heavily on local evidence.

This adaptive balance is what makes the LCAR formulation particularly robust across heterogeneous spatial settings.

Nota que aquí la media tiende al promedio de los vecinos pero no lo iguala completamente, y que en la sd, aunque disminuye a medida que aumenta el número de vecinos, permite más variabilidad.

## Interpretation of $\lambda$

Rather than being fixed by the analyst,  $\lambda$  is estimated from the data and acts as a **regulator**:

- When data are scarce or noisy, spatial borrowing stabilizes estimates.
- When information is sufficient, the model relies more on local evidence.

This balance is what makes the LCAR formulation particularly robust across heterogeneous spatial settings.