Rationale

Our objective is to identify outliers given a fitted simultaneous autoregressive (SAR) model. Consider n spatial units and p parameters, we derived the analytic expression to case influence analysis under the Gaussian assumption.

We then run through the Bayesian estimation for p parameters. It can be shown that $\boldsymbol{\beta}$ and σ^2 follow standard conjugacy result, but ρ needs extra care. The sampling algorithm draws $\boldsymbol{\beta}$ and σ^2 via Gibbs and ρ via rejection.

Coupling the analytic expression with r retained Markov chain Monte Carlo (MCMC) draws, we end up getting an $r \times n$ matrix. Each element calculates the case influence statistic given a set of MCMC parameters draw with single spatial observation removed.

PCA

Data

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}$$

Covariance

$$s_{j,k} = \frac{1}{n} \sum_{i=1}^{n} \left(x_{i,j} - \bar{x}_j\right)\left(x_{i,k} - \bar{x}_k\right) \quad j,k = 1,2,\ldots,p$$

Correlation

$$r_{j,k} = \frac{\sum_{i=1}^{n} \left(x_{i,j} - \bar{x}_j\right)\left(x_{i,k} - \bar{x}_k\right)}{n} \bigg/ s_j s_k$$

Eigendecomposition

$$\mathbf{A} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}$$

Eigenvalue

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix}$$

Eigenvector (`rotation` in `prcomp` and `loadings` in `princomp`)

$$\mathbf{V} = \begin{bmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,p} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ v_{p,1} & v_{p,2} & \cdots & v_{p,p} \end{bmatrix}$$

Reduced data (`x` in `prcomp` and `scores` in `princomp`)

$$\widetilde{\mathbf{X}} = \mathbf{X}\mathbf{V}$$

sPCA

Spatial data

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Spatial weight

$$\mathbf{W} = \begin{bmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,n} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n,1} & w_{n,2} & \cdots & w_{n,n} \end{bmatrix}$$

Spatial correlation (I and $LISA$)

$$\begin{aligned} I &= \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}\left(x_i - \bar{x}\right)\left(x_j - \bar{x}\right)}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}} \Bigg/ s^2 \\ &= \sum_{i=1}^{n} \frac{I_i}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}} \Bigg/ s^2 \end{aligned}$$

Spatial filtering