# Spatial Econometrics

## Lecture 2: Spatial weight matrix $\boldsymbol{W}$

**Andrzej Torój**

Institute of Econometrics – Department of Spatial Econometrics

# Outline

## Plan prezentacji

# Identification problem in spatial models (1)

- Consider the vector $\boldsymbol{y} = (y_1, ..., y_N)$ of spatially dependent observations.
- The spatial interdependence could potentially be formalized as:
  $y_1 = c_{1,1} + \alpha_{1,2} y_2 + \alpha_{1,3} y_3 + ... + \alpha_{1,N} y_N$
  $y_2 = c_{2,1} + \alpha_{2,1} y_1 + \alpha_{2,3} y_3 + ... + \alpha_{2,N} y_N$
  ...
  $y_N = c_{N,1} + \alpha_{N,1} y_1 + \alpha_{N,2} y_2 + ... + \alpha_{N,N-1} y_{N-1}$
- If all the elements of $\alpha_{i,j}$ were to be estimated (excluding constants), the number of estimated parameters would equal $N^2 - N$ and the number of observations $N$. Obviously impossible to do.

# Identification problem in spatial models (1)

- Consider the vector $\boldsymbol{y} = (y_1, ..., y_N)$ of spatially dependent observations.
- The spatial interdependence could potentially be formalized as:
  $y_1 = c_{1,1} + \alpha_{1,2} y_2 + \alpha_{1,3} y_3 + ... + \alpha_{1,N} y_N$
  $y_2 = c_{2,1} + \alpha_{2,1} y_1 + \alpha_{2,3} y_3 + ... + \alpha_{2,N} y_N$
  ...
  $y_N = c_{N,1} + \alpha_{N,1} y_1 + \alpha_{N,2} y_2 + ... + \alpha_{N,N-1} y_{N-1}$
- If all the elements of $\alpha_{i,j}$ were to be estimated (excluding constants), the number of estimated parameters would equal $N^2 - N$ and the number of observations $N$. Obviously impossible to do.

# Identification problem in spatial models (1)

- Consider the vector $\boldsymbol{y} = (y_1, ..., y_N)$ of spatially dependent observations.
- The spatial interdependence could potentially be formalized as:
  $y_1 = c_{1,1} + \alpha_{1,2} y_2 + \alpha_{1,3} y_3 + ... + \alpha_{1,N} y_N$
  $y_2 = c_{2,1} + \alpha_{2,1} y_1 + \alpha_{2,3} y_3 + ... + \alpha_{2,N} y_N$
  ...
  $y_N = c_{N,1} + \alpha_{N,1} y_1 + \alpha_{N,2} y_2 + ... + \alpha_{N,N-1} y_{N-1}$
- If all the elements of $\alpha_{i,j}$ were to be estimated (excluding constants), the number of estimated parameters would equal $N^2 - N$ and the number of observations $N$. Obviously impossible to do.

# Identification problem in spatial models (2)

- In practice, we make things feasible by treating the square matrix $\boldsymbol{\alpha} \equiv [\alpha_{i,j}]$ (with zero diagonal elements) as a product:

$$\boldsymbol{\alpha} = \rho \cdot \boldsymbol{W}$$

  - of unknown parameter $\rho$
  - of known, exogenous matrix $\boldsymbol{W}$
- Interpretation:
  - Non-zero elements of $i$-th row indicate regions linked to the $i$-th region, while proportions between positive values indicate the relative strength of this link.
  - $\rho$ is estimated. The exact interpretation depends on the construction details of $\boldsymbol{W}$ (see further), but if $\rho$ is not significantly different from 0, no spatial effects occur (and e.g. the classical linear regression model applies).

# Identification problem in spatial models (2)

- In practice, we make things feasible by treating the square matrix $\boldsymbol{\alpha} \equiv [\alpha_{i,j}]$ (with zero diagonal elements) as a product:

$$\boldsymbol{\alpha} = \rho \cdot \boldsymbol{W}$$

  - of unknown parameter $\rho$
  - of known, exogenous matrix $\boldsymbol{W}$
- Interpretation:
  - Non-zero elements of $i$-th row indicate regions linked to the $i$-th region, while proportions between positive values indicate the relative strength of this link.
  - $\rho$ is estimated. The exact interpretation depends on the construction details of $\boldsymbol{W}$ (see further), but if $\rho$ is not significantly different from 0, no spatial effects occur (and e.g. the classical linear regression model applies).

# Spatial weight matrix

- **W** is a central notion to spatial econometrics. It describes the network of relationships between any pair of units (e.g. regions).
  - Every element represents a link between a pair of regions.
- **W** is a square matrix. Its $i$-th row shall be interpreted as a vector of weights that define the impact of other regions on the $i$-th region.
- Matrix **W** has zero diagonal. I.e., the region does not impact on itself directly.
  - But it does indirectly: we impact the neighbour, and the induced change impact.

# Spatial weight matrix

- **W** is a central notion to spatial econometrics. It describes the network of relationships between any pair of units (e.g. regions).
  - Every element represents a link between a pair of regions.
- **W** is a square matrix. Its $i$-th row shall be interpreted as a vector of weights that define the impact of other regions on the $i$-th region.
- Matrix **W** has zero diagonal. I.e., the region does not impact on itself directly.
  - But it does indirectly: we impact the neighbour, and the induced change impact.

# Spatial weight matrix

- **W** is a central notion to spatial econometrics. It describes the network of relationships between any pair of units (e.g. regions).
  - Every element represents a link between a pair of regions.
- **W** is a square matrix. Its *i*-th row shall be interpreted as a vector of weights that define the impact of other regions on the *i*-th region.
- Matrix **W** has zero diagonal. I.e., the region does not impact on itself directly.
  - But it does indirectly: we impact the neighbour, and the induced change impact.

# Example matrix W

$$W = \begin{bmatrix}
0 & 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.33 & 0 & 0.33 & 0 & 0 & 0.33 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.2 & 0.2 & 0 & 0.2 & 0.2 & 0.2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.33 & 0.33 & 0 & 0 & 0.33 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.2 & 0 & 0.2 & 0 & 0 & 0.2 & 0 & 0.2 & 0.2 & 0 \\
0 & 0 & 0 & 0 & 0.2 & 0.2 & 0.2 & 0 & 0.2 & 0 & 0.2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.2 & 0.2 & 0.2 & 0.2 & 0 & 0.2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}$$

*Source:* ***Arbia (2014)***.

# How to build W matrix?

1. Define "space" adequately to the problem.
   1. geography?
   2. something else? (proximity / remoteness as a function of economic, social, cultural factors...)
2. Use an adequate measure for this definition.
   1. binary metrics: neighbourhood / adherence to a group
   2. continuous metrics: function of distance / "distance"
   3. hybrid metrics: choose $k$ nearest / "nearest" neighbours
3. Normalise matrix $\boldsymbol{W}$.

# How to build W matrix?

1. Define "space" adequately to the problem.
   1. geography?
   2. something else? (proximity / remoteness as a function of economic, social, cultural factors...)
2. Use an adequate measure for this definition.
   1. binary metrics: neighbourhood / adherence to a group
   2. continuous metrics: function of distance / "distance"
   3. hybrid metrics: choose $k$ nearest / "nearest" neighbours
3. Normalise matrix $\boldsymbol{W}$.

# How to build W matrix?

1. Define "space" adequately to the problem.
   1. geography?
   2. something else? (proximity / remoteness as a function of economic, social, cultural factors...)
2. Use an adequate measure for this definition.
   1. binary metrics: neighbourhood / adherence to a group
   2. continuous metrics: function of distance / "distance"
   3. hybrid metrics: choose $k$ nearest / "nearest" neighbours
3. Normalise matrix $\boldsymbol{W}$.

# Plan prezentacji

# Neighbourhood-based matrix

- Neighbourhood matrix is a special, most popular case.
- Regions (columns) neighbouring the $i$-th region are indicated as 1 in the $i$-th row (otherwise 0).
- The case of isolated regions might be problematic (Northern Ireland in UK, Sicilia in IT, etc.). One can attempt to avoid this by e.g. indicating the nearest region as the only neighbour, or one can correct the existing methods accordingly.
- Method adequate for polygons / regions (but not for points).

# Neighbourhood-based matrix

- Neighbourhood matrix is a special, most popular case.
- Regions (columns) neighbouring the $i$-th region are indicated as 1 in the $i$-th row (otherwise 0).
- The case of isolated regions might be problematic (Northern Ireland in UK, Sicilia in IT, etc.). One can attempt to avoid this by e.g. indicating the nearest region as the only neighbour, or one can correct the existing methods accordingly.
- Method adequate for polygons / regions (but not for points).

# Neighbourhood-based matrix

- Neighbourhood matrix is a special, most popular case.
- Regions (columns) neighbouring the $i$-th region are indicated as 1 in the $i$-th row (otherwise 0).
- The case of isolated regions might be problematic (Northern Ireland in UK, Sicilia in IT, etc.). One can attempt to avoid this by e.g. indicating the nearest region as the only neighbour, or one can correct the existing methods accordingly.
- Method adequate for polygons / regions (but not for points).

# Neighbourhood-based matrix

- Neighbourhood matrix is a special, most popular case.
- Regions (columns) neighbouring the $i$-th region are indicated as 1 in the $i$-th row (otherwise 0).
- The case of isolated regions might be problematic (Northern Ireland in UK, Sicilia in IT, etc.). One can attempt to avoid this by e.g. indicating the nearest region as the only neighbour, or one can correct the existing methods accordingly.
- Method adequate for polygons / regions (but not for points).

# Neighbourhood-based matrix: example

| Does... | ...neighbour... | ...? |
|---|---|---|
| USA | Canada | YES |
| Canada | USA | YES |
| USA | Mexico | YES |
| Mexico | USA | YES |
| Canada | Mexico | NO |
| Mexico | Canada | NO |

$$\boldsymbol{W}^S = \begin{array}{c} \begin{array}{ccc} US & CA & MX \end{array} \\ \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array} \qquad \boldsymbol{W} = \begin{array}{c} \begin{array}{ccc} US & CA & MX \end{array} \\ \begin{bmatrix} 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{array}$$

# Construction of matrix **W** in R based on neighbourhood

1. Create `SpatialPolygonDataFrame` (see: lecture 1). The map implies information about sharing borders by individual regions.
2. Create an object storing lists of neighbours – **nb**.
   ```
   contnb <- poly2nb(spatial_data, queen = T)
   ```
   - `queen = T` – treat as neighbours even when the shared border is only one vertex (chess players know where this comes from...)
3. Normalise by rows, creating a **listw** object.
   ```
   W_list <- nb2listw(contnb, style = "W")
   ```
   - If we have (and intend to keep) isolated regions, then the command `nb2listw` must be supplemented by the argument `zero.policy = TRUE`.
   - listw is an efficient form of storing the (usually sparse) matrix $\mathbf{W}$.

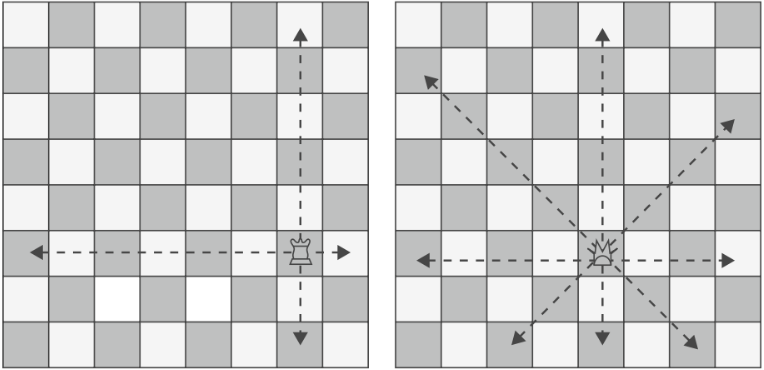
# Construction of matrix **W** in R based on neighbourhood

1. Create `SpatialPolygonDataFrame` (see: lecture 1). The map implies information about sharing borders by individual regions.
2. Create an object storing lists of neighbours – **nb**.
   `contnb <- poly2nb(spatial_data, queen = T)`
   - `queen = T` – treat as neighbours even when the shared border is only one vertex (chess players know where this comes from...)
3. Normalise by rows, creating a **listw** object.
   `W_list <- nb2listw(contnb, style = "W")`
   - If we have (and intend to keep) isolated regions, then the command `nb2listw` must be supplemented by the argument `zero.policy = TRUE`.
   - listw is an efficient form of storing the (usually sparse) matrix $\boldsymbol{W}$.

# Construction of matrix **W** in R based on neighbourhood

1. Create `SpatialPolygonDataFrame` (see: lecture 1). The map implies information about sharing borders by individual regions.
2. Create an object storing lists of neighbours – **nb**.
   `contnb <- poly2nb(spatial_data, queen = T)`
   - `queen = T` – treat as neighbours even when the shared border is only one vertex (chess players know where this comes from...)
3. Normalise by rows, creating a **listw** object.
   `W_list <- nb2listw(contnb, style = "W")`
   - If we have (and intend to keep) isolated regions, then the command `nb2listw` must be supplemented by the argument `zero.policy = TRUE`.
   - listw is an efficient form of storing the (usually sparse) matrix $\boldsymbol{W}$.

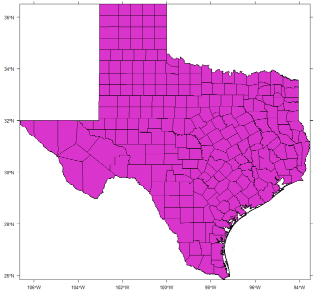
# Queen criterion (queen = T) (1)



Source: *Arbia (2014)*.

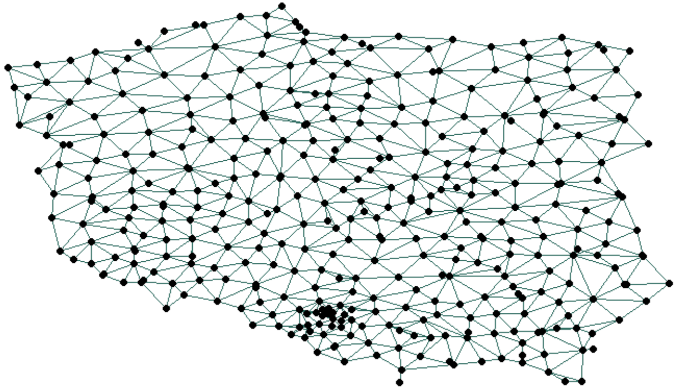
# Queen criterion (queen = T) (2)

Appears as a virtual problem when we think about Polish poviats. But real if we look at counties in Texas (US) or point data aggregated into "regions" on a rectangular grid (we'll come back to it).



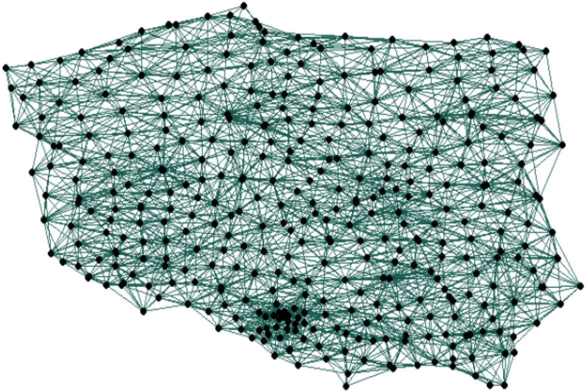
# Neighbourhood of order 1 for Polish poviats

Function: `plot.nb`



# Neighbourhood of order 1 and 2 for Polish poviats

Every neighbour of my neighbour is also my neighbour (`nblag` and `nblag_cumul`)



# Construction of distance-based matrix

- A alternative way of building matrix $\boldsymbol{W}$ is allowing for direct interactions between all regions. The intensity of link depends on the distance.
- Let $d_{i,j}$ denote the distance between $i$ and $j$. Then:

$$\boldsymbol{W} = \begin{bmatrix} 0 & \frac{1}{(d_{1,2})^\gamma} & \cdots & \frac{1}{\left(d_{1,N}\right)^\gamma} \\ \frac{1}{(d_{1,2})^\gamma} & 0 & \cdots & \frac{1}{\left(d_{2,N}\right)^\gamma} \\ \vdots & \vdots & \ddots & \\ \frac{1}{\left(d_{1,N}\right)^\gamma} & \frac{1}{\left(d_{2,N}\right)^\gamma} & & 0 \end{bmatrix}$$

- $\gamma$ is a calibrated parameter of decay. In the absence of any other evidence, one could assume that e.g. $\gamma = 1$.

# Construction of distance-based matrix

- A alternative way of building matrix $\boldsymbol{W}$ is allowing for direct interactions between all regions. The intensity of link depends on the distance.
- Let $d_{i,j}$ denote the distance between $i$ and $j$. Then:

$$\boldsymbol{W} = \begin{bmatrix} 0 & \frac{1}{(d_{1,2})^{\gamma}} & \cdots & \frac{1}{\left(d_{1,N}\right)^{\gamma}} \\ \frac{1}{(d_{1,2})^{\gamma}} & 0 & \cdots & \frac{1}{\left(d_{2,N}\right)^{\gamma}} \\ \vdots & \vdots & \ddots & \\ \frac{1}{\left(d_{1,N}\right)^{\gamma}} & \frac{1}{\left(d_{2,N}\right)^{\gamma}} & & 0 \end{bmatrix}$$

- $\gamma$ is a calibrated parameter of decay. In the absence of any other evidence, one could assume that e.g. $\gamma = 1$.

# Construction of distance-based matrix

- A alternative way of building matrix $\boldsymbol{W}$ is allowing for direct interactions between all regions. The intensity of link depends on the distance.
- Let $d_{i,j}$ denote the distance between $i$ and $j$. Then:

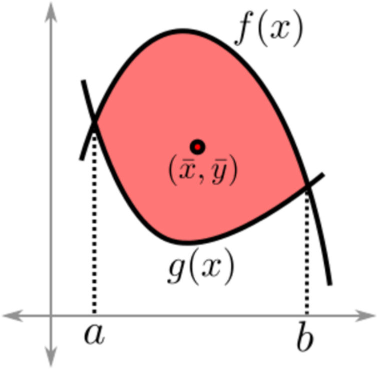
$$\boldsymbol{W} = \begin{bmatrix} 0 & \frac{1}{(d_{1,2})^{\gamma}} & \cdots & \frac{1}{\left(d_{1,N}\right)^{\gamma}} \\ \frac{1}{(d_{1,2})^{\gamma}} & 0 & \cdots & \frac{1}{\left(d_{2,N}\right)^{\gamma}} \\ \vdots & \vdots & \ddots & \\ \frac{1}{\left(d_{1,N}\right)^{\gamma}} & \frac{1}{\left(d_{2,N}\right)^{\gamma}} & & 0 \end{bmatrix}$$

- $\gamma$ is a calibrated parameter of decay. In the absence of any other evidence, one could assume that e.g. $\gamma = 1$.

# Construction of distance-based matrix in R

- The function `DistanceMatrix` is required.
- It takes an object of type `SpatialPolygonDataFrame` as an argument.
- The function works correctly only if points / vertices of polygons are supplied as degrees of longitude and latitude.
    - In the maps from CODGiK, this is not the case.
    - This is why we need to `spTransform`.
    - When working with your own data, one should always look up the order of coordinates' magnitude or, if available, the coding standard of coordinates.
- The distances are computed either between pairs of points (for a map of points) or pairs of **centroids** (for a map of polygons).

# Centroids -- geometric centers of gravity (1)

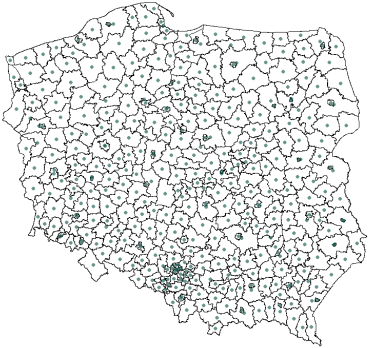



$$\bar{x} = \frac{\int_a^b \int_{g(x)}^{f(x)} x dy dx}{\int_a^b \int_{g(x)}^{f(x)} dy dx} = \frac{\int_a^b x(f(x) - g(x)) dx}{\int_a^b (f(x) - g(x)) dx}.$$

$$\bar{y} = \frac{\int_a^b \int_{g(x)}^{f(x)} y\, dy\, dx}{\int_a^b \int_{g(x)}^{f(x)} dy\, dx} = \frac{\int_a^b \frac{1}{2}(f(x)^2 - g(x)^2)\, dx}{\int_a^b (f(x) - g(x))\, dx}.$$

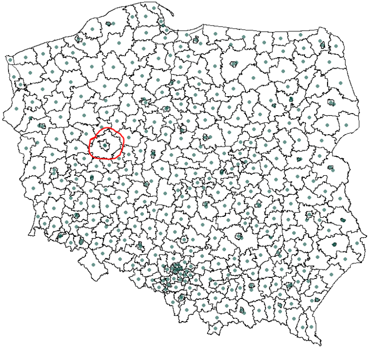
In each case, the denominator is the area of the figure.
The real-world borders of regions are usually not represented as functional forms :-) (Texas might be an exception...), so R is using numerical approximations to the above integrals.

## Centroids (2)



Paradox – notice the location of centroid for the *poznański* poviat (not the same as poviat *city Poznań*!).

# Centroids (3)



Paradox – notice the location of centroid for the *poznański* poviat (not the same as poviat *city Poznań*!).

# Plan prezentacji

# Neighbour selection by distance criterion

- Elements of matrix $\boldsymbol{W}$ are put equal to 1, if the distance between two regions does not exceed a pre-specified value, or 0 otherwise.
- The method works well, when:
  - our data does not represent spatial polygons, but spatial points (and the logic of sharing borders does not apply);
  - we intend to avoid the presence of isolated regions.
- Instead of `poly2nb` we use the function `dnearneigh` (after obtaining the centroids via function `coordinates`).

# k nearest neighbours method

- Elements of matrix $\boldsymbol{W} = [w_{i,j}]$ are equal to 1, if region $j$ belongs to $k$ geographically nearest regions for $i$ (0 otherwise).
    - Numerical investigations confirm that modelling results are usually robust to the choice of $k$ (***LeSage and Pace, 2014***).
- The method works well, when:
    - our data does not represent spatial polygons, but spatial points (and the logic of sharing borders does not apply);
    - we intend to avoid the presence of isolated regions.
- Instead of `poly2nb` we use the function `knearneigh` (on centroids, again).

# Figurative measures of distance

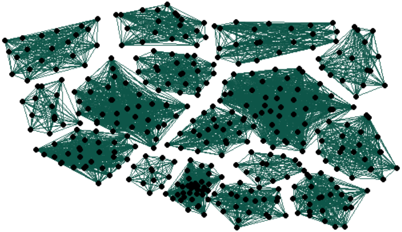
- Sometimes we measure distance in a non-geographical, figurative way (*Corrado, Fingleton, 2012*).
  - Instead, we look at cultural, social or economic proximity of regions – not necessarily the neighbouring or nearby ones. E.g., in some respects, Warsaw is more connected to Poznań than to Otwock.
- Then we have to construct $\boldsymbol{W}$ on our own (example provided in the R file).

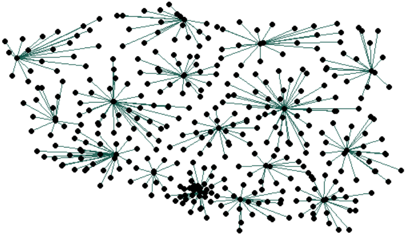
**But let's be careful!**

Using such an approach, we are more exposed to the risk of breaking the fundamental assumption: about exogeneity of $\boldsymbol{W}$. As a result, the estimation of spatial models could be **inconsistent**.
Solution: network co-evolution models (cf. *Franzese, Hays, Kachi 2012*).

# User-supplied neighbourhood matrices

- For example: neighbourhood only within one voivodship or only capitals of voivodships being neighbours.





# Plan prezentacji

# Why do we normalise W?

- To obtain an intuitive interpretation of $\rho$ as the weighted average impact of neighbours.
- To avoid the risk of inverting a near-singular matrix.
- To avoid numerical issues related to different scaling of variables.
- To avoid "explosive" spatial multipliers implied by $\rho$ (by analogy to time series econometrics, where autoregression parameter is expected to be $< 1$ in modulus – we'll come back to that when discussing spatial panels).

# Why do we normalise W?

- To obtain an intuitive interpretation of $\rho$ as the weighted average impact of neighbours.
- To avoid the risk of inverting a near-singular matrix.
- To avoid numerical issues related to different scaling of variables.
- To avoid "explosive" spatial multipliers implied by $\rho$ (by analogy to time series econometrics, where autoregression parameter is expected to be $< 1$ in modulus – we'll come back to that when discussing spatial panels).

# Why do we normalise W?

- To obtain an intuitive interpretation of $\rho$ as the weighted average impact of neighbours.
- To avoid the risk of inverting a near-singular matrix.
- To avoid numerical issues related to different scaling of variables.
- To avoid "explosive" spatial multipliers implied by $\rho$ (by analogy to time series econometrics, where autoregression parameter is expected to be $< 1$ in modulus – we'll come back to that when discussing spatial panels).

# Why do we normalise W?

- To obtain an intuitive interpretation of $\rho$ as the weighted average impact of neighbours.
- To avoid the risk of inverting a near-singular matrix.
- To avoid numerical issues related to different scaling of variables.
- To avoid "explosive" spatial multipliers implied by $\rho$ (by analogy to time series econometrics, where autoregression parameter is expected to be $< 1$ in modulus – we'll come back to that when discussing spatial panels).

# Normalisation techniques (1)

- Most frequently, we normalise $\boldsymbol{W}$ row by row, so as to make the elements of each row sum to unity (*row-stochastic*, *row-standardized*).
  - `nb2listw(..., style = "W")`
  - side effect: asymmetry of normalised **W** (against theoretical arguments: ***Vega and Elhorst, 2014***)
- Instead of "W" we can also use:
  - **"B"**: not normalised
  - **"C"**: normalised by scalar (all elements sum to $N$, but individual rows not necessarily sum to 1)
  - **"N"**: normalised by scalar (all elements sum to 1)
  - **"S"**: ***Tiefelsdorf, Griffith, Boots (1999)***
  - **"minmax"**: ***Kelejian, Prucha (2010)***

# Normalisation by row – always good?

- **Advantage**: vector **Wx** (spatial lag – more about it in the next lecture) can be interpreted as a weighted average of the variables for the units connected to the first, second, third unit... etc. (e.g. for their neighbours).
- **Disadvantage**: whose apartment is louder? 's or 's?

| | | | |
|---|---|---|---|
| 0 dB |  | **150 dB** |  |
| 0 dB | 0 dB | 0 dB | 0 dB |
| 0 dB | 0 dB | 0 dB | 0 dB |

When elements of neighbourhood-based**W** (queen-criterion) are normalised by row:

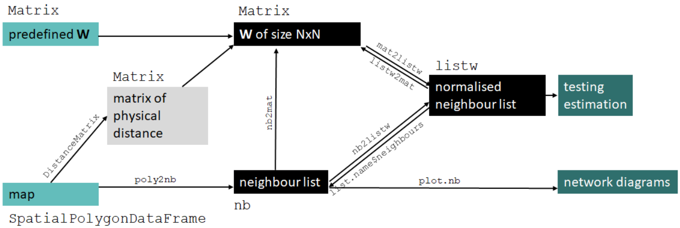
- : $\frac{1}{5}(0+0+0+0+150) = 30dB$ (which is more silent than...)
- : $\frac{1}{3}(0+0+150) = 50dB$ (is that intuitive...?)

# Normalisation techniques (2)

- Alternative method: normalise by scalar – maximum modulus of $\boldsymbol{W}$'s eigenvalue:
$$\mathbf{W}^* = \frac{\mathbf{W}}{\lambda_{max}}$$
- Most popular strategy among normalisations by scalar.
- Advantage: $\frac{1}{\lambda_{min}} < \rho < 1$.
- Disadvantage: $\rho$ not interpretable in a standard way.

# How does R store matrix W?

Depending on the input information and the end-point of the analysis, we use and convert different objects all along the way.



# How to select the optimum $\mathbf{W}$?

- Usually few arguments to support any choice (***Anselin, 2002***). Best guess.
- ***Neumayer, Plumper (2016)*** advise to let the theory guide you (sometimes difficult...).
- Post-estimation comparison of log-likelihood or Bayesian posterior model probabilities for different $\boldsymbol{W}$.

**The fundamental question to address:**

Does $\boldsymbol{W}$ adequately represent the network of direct links between units in the context of the analysed dependent variable?

- Consequences of a mistake:
  - low power of diagnostic tests;
  - inconsistency and bias of estimators;
  - weak statistical identification of models.

# Homework 2

For the spatial dataset considered in the first homework, build three additional **W** matrices:

1. neighbourhood (order 1), but normalised with the highest eigenvalue (tip: you can do it with the matrix immediately, and then transform it into **listw** without normalisation).
2. based on inverted squared distances, but only up to 200 km (above that – no link at all).
3. based on Euclidian distance between 3 standardised variables for pairs of regions $i, j$ $\left(\sqrt{(x_{1,i} - x_{1,j})^2 + (x_{2,i} - x_{2,j})^2 + (x_{3,i} - x_{3,j})^2}\right)$; these variables should, in our view, correctly reflect the network of connections related to the variable presented in Homework 1.

In future homeworks, one can consider other matrices than those built here.