



# ÉCONOMÉTRIE SPATIALE

Master 2 APE

Introduction et Chapitre 1

Gilles de Truchis

[gilles.detruchis@gmail.com](mailto:gilles.detruchis@gmail.com)

Site : [www.varenes-ecofin.com/](http://www.varenes-ecofin.com/)



# LES CHAPITRES DU COURS

- 1. Généralités
- 2. Introduction

- 3. Matrice de pondération
  - 3.1 Concepts
  - 3.2 Matrices

# GÉNÉRALITÉS

# ORGANISATION DU COURS

- ▶ Le cours est organisé en 6 séances de 3h + 1 séance de 2h
  
- ▶ Le cours s'articule autour de 5 chapitres + 1 annexe
  - C1 Matrices de pondérations spatiales
  - C2 Autocorrélation spatiale
  - C3 Modèles spatialement dépendants
  - C4 Modèles spatiaux avec hétérogénéité
  - A1 Annexe sur les techniques d'estimation
  - C5 Introduction à GEODA
  
- ▶ Objectif
  - ▶ Apporter un savoir technique pour l'analyse en économie spatiale
    - ▶ analyser des données spatiales
    - ▶ spécifier un modèle
    - ▶ estimer et tester ce modèle
  
- ▶ La première partie du cours sera théorique et la seconde appliquée
  - ▶ Études de cas sur les logiciels gratuits GeoDa / GeoDa Space / GWR4

# ÉVALUATION

- ▶ L'évaluation se fera sur la base d'un dossier (10 pages max)
  
- ▶ Sur le plan technique vous devrez :
  - ▶ Créer une base de donnée : les shapes et les données sont fournis
  - ▶ Mener à bien une étude économétrique rigoureuse
    - ▶ sélectionner les variables appropriées
    - ▶ implémenter les statistiques usuelles
    - ▶ spécifier un modèle
    - ▶ tester la robustesse de vos résultats
  - ▶ Détailler sur le plan formel les techniques utilisées
  
- ▶ Sur le plan de la problématique traitée vous devrez :
  - ▶ Replacer la problématique dans son contexte socio-économique
  - ▶ Commenter les résultats économétriques
  - ▶ Tirer des conclusions en termes de politique économique
  - ▶ Conclure par une ouverture
  
- ▶ Lors de la dernière séance chaque binôme (trinôme) présentera oralement son dossier (environ 15 min + 5 min questions)

# INTRODUCTION

# LE CADRE DE L'ÉCONOMÉTRIE SPATIALE

- ▶ L'économétrie spatiale s'inscrit dans un cadre bien plus large
  - ⇒ l'analyse spatiale
  
- ▶ Outils s'appliquant à de nombreux domaines
  - ⇒ géographie
  - ⇒ exploitation minière
  - ⇒ épidémiologie
  - ⇒ agronomie
  - ⇒ traitement d'image
  - ⇒ foresterie
  - ⇒ ...

# LES APPLICATIONS DE L'ÉCONOMÉTRIE SPATIALE

- ▶ La statistique spatiale a de très vastes applications
- ▶ Dans le cadre spécifique de l'économétrie les applications sont également nombreuses
  - ▶ Economie urbaine
  - ▶ Economie internationale
  - ▶ Economie publique
  - ▶ Analyse de la croissance
  - ▶ ...

# LE DÉVELOPPEMENT DE L'ÉCONOMÉTRIE SPATIALE

- ▶ L'économétrie spatiale connaît un fort développement
- ⇒ croissance massive des données géo-référencées
  
- ▶ Émergence des **systèmes d'information géographiques (SIG)** caractérisés par
  - ⇒ un logiciel
  - ⇒ des données géographiques
  - ⇒ une capacité de traitement
  - ⇒ des toolbox permettant une utilisation simple

# LOGICIEL ET DONNÉES

- ▶ L'aspect empirique du cours s'appuiera sur le logiciel GeoDa
- ▶ GeoDa est disponible gratuitement sur internet  
⇒ <https://geodacenter.asu.edu/software/downloads>
- ▶ Les données utilisées seront disponibles sur **cours en ligne** ou sur mon site  
⇒ [www.varennes-ecofin.com/](http://www.varennes-ecofin.com/)

# CADRE FORMEL DE L'ÉCONOMÉTRIE SPATIALE

- ▶ On cherche à étudier des observations géographiques
- ▶ Ces observations sont des processus aléatoires
- ▶ On définit  $X = \{x_s, s \in S\}$  un processus aléatoire avec
  - ▶  $S$  un ensemble spatial
  - ▶  $s$  un site d'observation
  - ▶  $x_s$  un événement de  $\Omega$ , l'ensemble des états
- ▶  $s \in S$  est soit déterministe, soit aléatoire

## CADRE FORMEL DE L'ÉCONOMÉTRIE SPATIALE

- ▶ L'ensemble spatial est
  - ▶ unidimensionnel si  $S \subseteq \mathbb{R}$ 
    - ▶ coordonnées de type  $s = x$   
⇒ e.g. chromatographie
  - ▶ bidimensionnel si  $S \subseteq \mathbb{R}^2$ 
    - ▶ coordonnées latitude/longitude de type  $s = (x, y)$   
⇒ e.g. agronomie
  - ▶ tridimensionnel si  $S \subseteq \mathbb{R}^3$ 
    - ▶ coordonnées latitude/longitude/altitude de type  $s = (x, y, z)$   
⇒ e.g. prospection minière
  - ▶ dynamique si une dimension temporelle vient indexer l'observation  
⇒ e.g. migration

$$(s, t) \in \mathbb{R}^2 \times \mathbb{R}^+$$

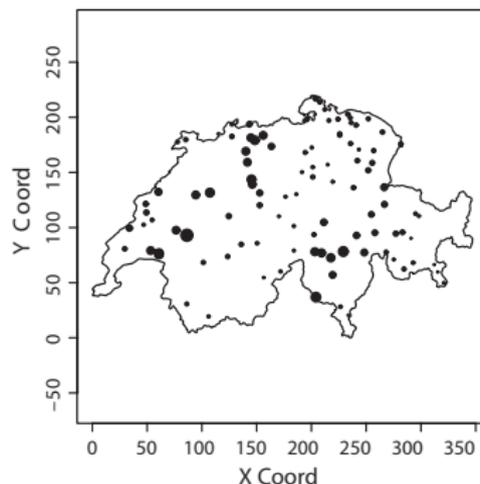
# TYPOLOGIE DES DONNÉES

## ► Données géostatistiques

### ► Pour $S \subseteq \mathbb{R}^d$

- on dispose d'un champ  $X = \{x_s, s \in S\}$
- observé en  $n$  sites fixés  $\{s_1, \dots, s_n\} \subset S$
- les sites ne sont pas forcément disposés régulièrement

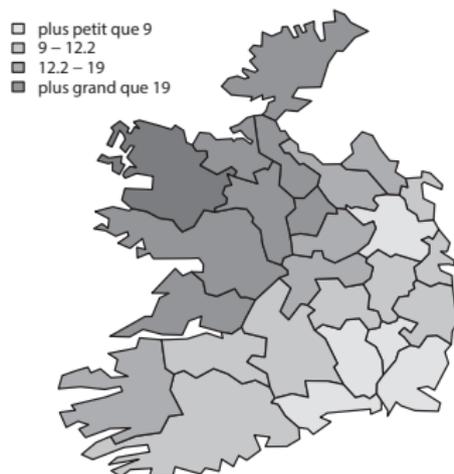
### ► Exemple : données pluviométriques Suisse au passage du nuage de Chernobyl



## TYPOLOGIE DES DONNÉES

## ▶ Données latticielles (sur réseau fixe)

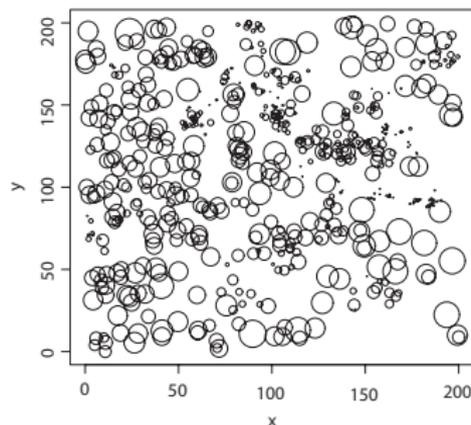
- ▶ Pour  $S \subseteq \mathbb{R}^d$  et un champ  $X = \{x_s, s \in S\}$ 
  - ▶  $S$  est discret et non-aléatoire (réseau fixe)
  - ▶ les sites  $s$  représentent des unités géographiques
  - ▶  $x_s$  intègre une quantité d'intérêt de cette unité  $s$

▶ Exemple : % d'individus du groupe sanguin  $A$  dans les comtés Irlandais

# TYPOLOGIE DES DONNÉES

## ► Données ponctuelles

- Pour  $S \subset \mathbb{R}^d$  et un champ  $X = \{x_s, s \in S\}$ 
  - l'ensemble des sites  $y = \{y_1, \dots, y_n\}, y_i \in S$  est aléatoire
  - avec  $n$  le nombre de sites aléatoires également
  - avec  $x$  une réalisation d'un processus ponctuel  $X$
  
- Exemple : longueur des aiguilles de pins observées en  $y$
- ⇒ l'exemple pose la question de la répartition des points



# CHAMP ALÉATOIRE SPATIALE

- ▶ En séries temporelles, la dépendance est un phénomène fréquent
  - ⇒ e.g. l'observation aujourd'hui dépend de l'observation d'hier
- ▶ En statistique spatiale, la dépendance est également fréquente
  - ⇒ e.g. l'observation en un site dépend des observations avoisinantes
- ▶ Importance du phénomène de dépendance
  - ▶ la dépendance crée une redondance d'information
  - ▶ cette information est exploitable pour la prévision
  - ⇒ néanmoins, la dépendance **modifie** les comportements **statistiques usuels**

# NON-CAUSALITÉ SPATIALE

- ▶ En séries temporelles, les processus sont généralement **causaux**
  - ▶ Le temps est orienté du passé vers le futur
- ▶ En statistique spatiale, les processus sont généralement **non-causaux**
- ▶ Exemple : on s'intéresse à la rentabilité des boulangeries à Paris
  - ▶ La dynamique de rentabilité de  $b_{s,t}$  pourrait dépendre de  $b_{s,t-1}$ 
    - ⇒ le processus est causal car il regarde uniquement vers le passé
  - ▶ La rentabilité de  $b_{s,t}$  pourrait dépendre  $b_{s+1,t}$  et  $b_{s-1,t}$ 
    - ⇒ le processus est non-causal car il regarde (e.g.) à gauche et à droite

# L'APPROCHE GÉOSTATISTIQUE

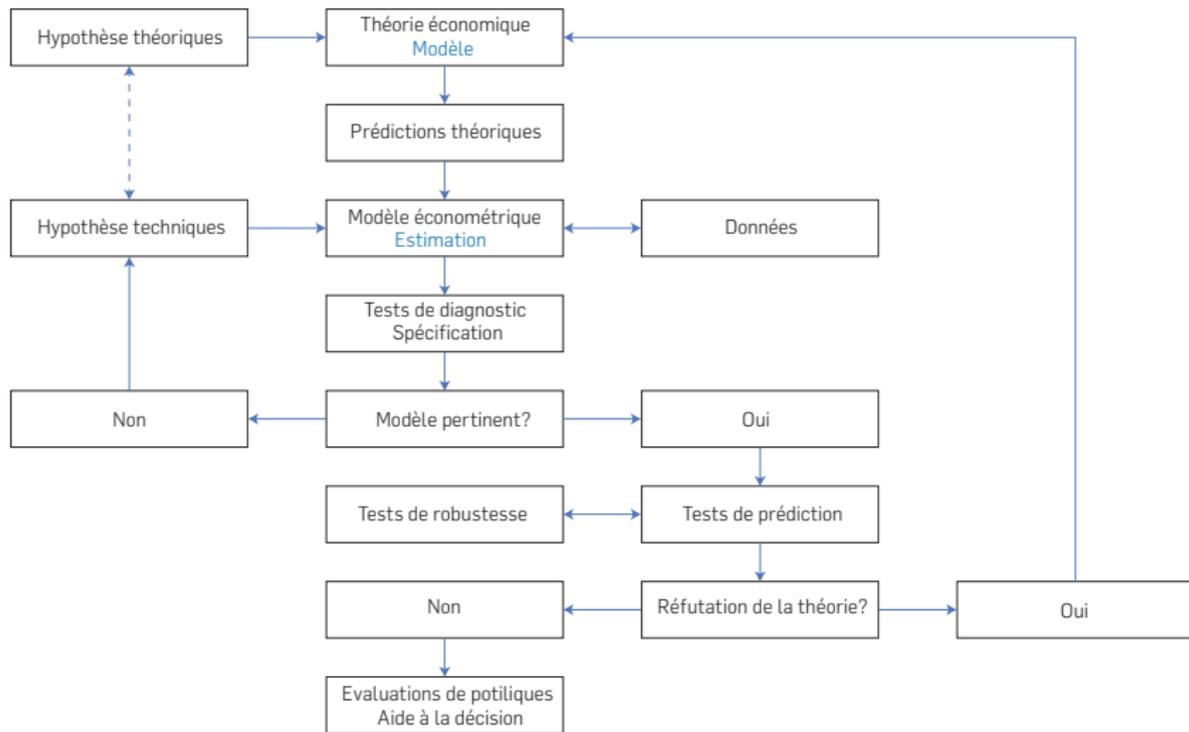
Deux grandes approches sont possibles

- ▶ L'approche géostatistique
  - ▶ repose sur des données géostatistique
  - ⇒  $S$  est continu sur  $\mathbb{R}^d$
  
  - ▶  $X$  est modélisé au **second ordre**
  - ⇒ on se base sur la fonction de **covariance** de  $X$
  - ⇒ on se base également sur le **variogramme** de  $X$

# L'APPROCHE AUTOREGRESSIVE

- ▶ L'approche autoregressive (AR)
  - ▶ repose sur des données latticielles
  - ⇒ En économétrie, le réseau sera souvent irrégulier
  
  - ▶ L'analyse de  $X$  sera basée sur un modèle
  - ⇒ ces modèles permettent de tenir compte de la dépendance spatiale
  - ⇒ le modèle AR retenu conditionnera la structure de **dépendance**

## MÉTHODOLOGIE ÉCONOMÉTRIQUE



# LES MODÈLES ÉCONOMÉTRIQUES

- ▶ Modèle en coupe transversale

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

- ▶ Hypothèses standards :  $E(\varepsilon_i) = 0$ ;  $E(\varepsilon_i^2) = \sigma_\varepsilon^2$ ;  $E(\varepsilon_i \varepsilon_j) = 0$

- ▶ Modèles de séries temporelles

$$y_t = \alpha + \beta x_t + \varepsilon_t; \quad y_t = \alpha + \rho y_{t-1} + \beta x_t + \varepsilon_t$$

- ▶ Hypothèses standards :  $E(\varepsilon_t) = 0$ ;  $E(\varepsilon_t^2) = \sigma_\varepsilon^2$ ;  $E(\varepsilon_t \varepsilon_s) = 0$

- ▶ Modèles de panel

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}; \quad y_{it} = \alpha + \rho y_{it-1} + \beta x_{it} + \varepsilon_t$$

- ▶ Hypothèses standards :  $\varepsilon_{it} \sim \text{i. i. d. } (0, \sigma_\varepsilon^2)$

# MODÈLES SPATIAUX

## ► Modèle spatial

$$y_i = \alpha + \rho \sum_{j=1}^n w_{ij} y_j + \beta x_i + \varepsilon_i$$

- Hypothèses standards :  $\varepsilon_i \sim \text{i. i. d. } (0, \sigma_\varepsilon^2)$
- ⇒ A la différence des modèles en coupe, les individus peuvent ici interagir
- ⇒ Le schéma d'interaction est déterminé par la matrice  $W$

## ► Modèles de panel spatial

$$y_{it} = \alpha_i + \rho \sum_{j=1}^n w_{ij} y_{jt} + \beta x_{it} + \varepsilon_{it}$$

- Hypothèses standards :  $\varepsilon_{it} \sim \text{i. i. d. } (0, \sigma_\varepsilon^2)$

# MATRICE DE PONDÉRATION

# CONTENTS

- 1. Généralités
- 2. Introduction

- 3. Matrice de pondération
  - 3.1 Concepts
  - 3.2 Matrices

# CONCEPT DE PONDÉRATION

- ▶ Les interactions spatiales impliquent :  $Cov(x_i, x_j) \neq 0$
  - ▶ On peut formuler plusieurs hypothèses d'interaction
    - ▶ les localisations  $i$  et  $j$  interagissent car elles sont voisines
    - ▶ Le degré d'interaction diminue avec la distance
    - ▶ Le degré d'interaction augmente avec la taille de la frontière
  - ▶ Pour définir les voisins de chaque localisation on utilise une matrice
- ⇒ **matrice de pondération**

# CONCEPT DE CONTIGUÏTÉ

- ▶ Il existe différents types de contiguïté d'unités irrégulières ou régulières
  - ▶ e.g. contiguïté de type "tour" (rook) : frontières
  - ▶ e.g. contiguïté de type "reine" (queen) : sommets + frontières

Tour

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Fou

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

Reine

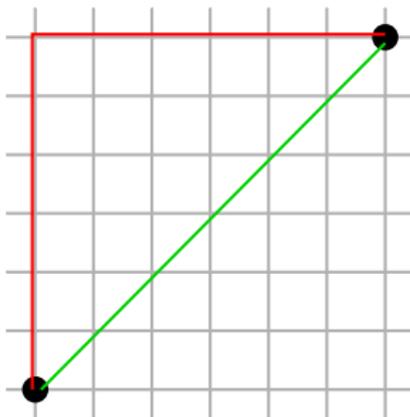
1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25

# CONCEPTS DE DISTANCE

- ▶ On nomme distance sur un ensemble  $E$ , une application  $d$  définie sur  $E^2$  et à valeurs dans l'ensemble  $\mathbb{R}^+$  vérifiant les propriétés suivantes
  - ▶ symétrie :  $\forall (a, b) \in E^2, \quad d(a, b) = d(b, a)$
  - ▶ séparation :  $\forall (a, b) \in E^2, \quad d(a, b) = 0 \iff a = b$
  - ▶ inégalité triangulaire :  $\forall (a, b, c) \in E^3, \quad d(a, c) \leq d(a, b) + d(b, c)$
  
- ▶ De cette définition découle différentes mesures pour  $a, b \in \mathbb{R}^n$ 
  - ▶ distance de Manhattan :  $d(a, b) = \sum_{i=1}^n |a_i - b_i|$
  - ▶ distance de Minkowski :  $d(a, b) = \sqrt[p]{\sum_{i=1}^n |a_i - b_i|^p}$
  - ▶ distance euclidienne :  $d(a, b) = \sqrt{\sum_{i=1}^n |a_i - b_i|^2}$
  - ▶ distance de Tchebychev :  $d(a, b) = \lim_{p \rightarrow \infty} \sqrt[p]{\sum_{i=1}^n |a_i - b_i|^p} = \max_i |a_i - b_i|$

## EXEMPLE DISTANCE

- ▶ Exemple si  $a, b \in \mathbb{R}^2$  :  $a = (x_a, y_a)$  et  $b = (x_b, y_b)$
- ▶ distance de Manhattan :  $d(a, b) = |a_1 - b_1| + |a_2 - b_2| = |x_a - x_b| + |y_a - y_b|$
- ▶ distance de euclidienne :  $d(a, b) = \sqrt{|x_a - x_b|^2 + |y_a - y_b|^2}$



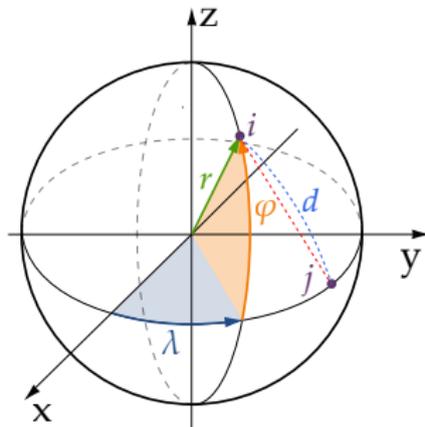
## DISTANCE ORTHODROMIQUE

- ▶ Soit  $r$  le rayon de la terre ( $\approx 6371\text{km}$ )
  - ▶ Soit  $\lambda_i^{(deg)}$  et  $\varphi_i^{(deg)}$  la longitude  $\in [0^\circ, 360^\circ]$  et la latitude  $\in [0^\circ, 90^\circ]$  d'un point  $i$
  - ▶ Soit  $\lambda_j^{(deg)}$  et  $\varphi_j^{(deg)}$  la longitude  $\in [0^\circ, 360^\circ]$  et la latitude  $\in [0^\circ, 90^\circ]$  d'un point  $j$
- ▶ La distance orthodromique entre  $i$  et  $j$  est alors donnée par

$$d_{ij} = 2r \arcsin \sqrt{\sin^2 \left( \frac{|\varphi_j - \varphi_i|}{2} \right) + \cos(\varphi_i) \cos(\varphi_j) \sin^2 \left( \frac{|\lambda_j - \lambda_i|}{2} \right)}$$

- ▶  $\lambda_i$  et  $\varphi_i$  sont des valeurs convertis en radian

$$\lambda_i = \lambda_i^{(rad)} = \lambda_i^{(deg)} \times \pi/180 \text{ et } \varphi_i = \varphi_i^{(rad)} = \varphi_i^{(deg)} \times \pi/180$$



# CONTENTS

- 1. Généralités
- 2. Introduction

- 3. Matrice de pondération
  - 3.1 Concepts
  - 3.2 Matrices

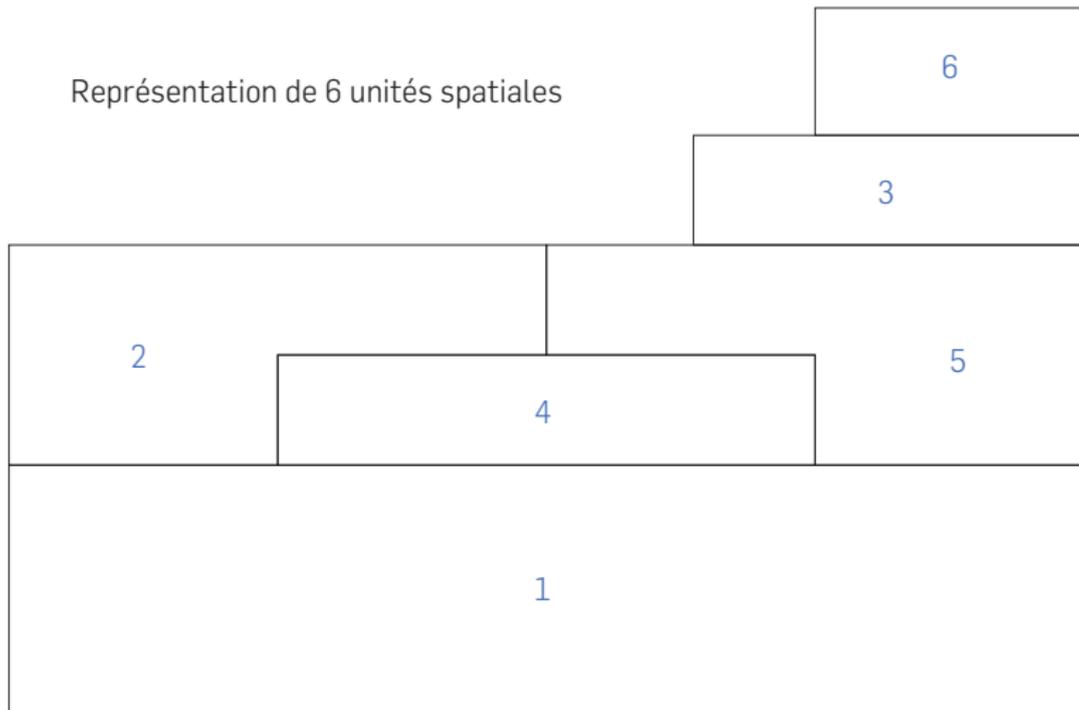
# MATRICE DE PONDÉRATION

- ▶ Il s'agit d'une matrice carrée de taille  $N \times N$  notée  $W$
- ▶ Les éléments  $w_{ij}$  de cette matrice sont généralement binaires
  - ▶ e.g.  $w_{ij} = 1$  si  $i$  et  $j$  sont contiguës, sinon  $w_{ij} = 0$
  - ▶ e.g.  $w_{ij} = 1$  si  $d_{ij}$  est inférieur à une bande de distance, sinon  $w_{ij} = 0$
  - ▶ e.g.  $w_{ij} = 1$  si  $i$  et  $j$  sont plus proche voisins, sinon  $w_{ij} = 0$
  - ▶ e.g. si  $i = j$ ,  $w_{ij} = 0$
- ▶ Il est possible de standardiser la matrice
  - ⇒ standardisation par la moyenne pondérée des valeurs voisines

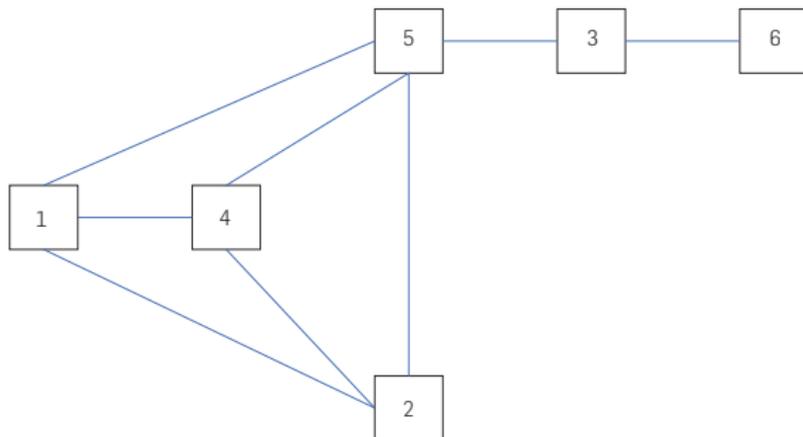
$$w_{ij}^s = w_{ij} \left( \sum_j w_{ij} \right)^{-1} \iff \sum_j w_{ij}^s = 1$$

## MATRICE DE PONDÉRATION : EXEMPLE

Représentation de 6 unités spatiales



## MATRICE DE CONTIGUÏTÉ : GRAPHE



- Déterminer la matrice  $W$  de contiguïté d'ordre 1

## MATRICE DE CONTIGUÏTÉ : EXEMPLE

- ▶ La matrice de contiguïté correspondante est donc

## MATRICE DE CONTIGUÏTÉ : EXEMPLE

- ▶ La matrice de contiguïté correspondante est donc

$$W = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

- ▶ Notons qu'il s'agit d'une matrice symétrique
- ▶ En standardisant on obtient ?

## MATRICE DE CONTIGUÏTÉ : EXEMPLE

- ▶ La matrice de contiguïté correspondante est donc

$$W = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

- ▶ Notons qu'il s'agit d'une matrice symétrique
- ▶ En standardisant on obtient ?

$$W^{(s)} = \begin{pmatrix} 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 1/3 & 0 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 0 & 0 & 1/3 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

MATRICE DE CONTIGUÏTÉ D'ORDRE  $k$ 

## ► Définition

- $i$  et  $j$  sont contigus à l'ordre  $p$  si  $p$  est le nombre minimale de frontières à traverser pour aller de  $i$  à  $j$

► **Erreur de construction** une matrice d'ordre  $k > 1$ 

- L'approche à **proscrire** et d'utiliser une puissance de  $W$  (e.g. avec  $k = 2$ )

$$W'W = \begin{pmatrix} 3 & 2 & 1 & 2 & 2 & 0 \\ 2 & 3 & 1 & 2 & 2 & 0 \\ 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 3 & 2 & 0 \\ 2 & 2 & 0 & 2 & 4 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

# MATRICE DE CONTIGUÏTÉ D'ORDRE $k$

## ► **Méthode de construction** d'une matrice d'ordre $k > 1$

- Utiliser une représentation en graphe du réseau
- Algorithme de Dijkstra modifié : détermine le nombre de "pas" jusqu'au plus proche voisin  
⇒ le nombre de "pas" définit l'ordre de contiguïté
- Par exemple, la matrice de contiguïté à l'ordre  $k = 2$

MATRICE DE CONTIGÜITÉ D'ORDRE  $k$ ► **Méthode de construction** d'une matrice d'ordre  $k > 1$ 

- Utiliser une représentation en graphe du réseau
- Algorithme de Dijkstra modifié : détermine le nombre de "pas" jusqu'au plus proche voisin  
⇒ le nombre de "pas" définit l'ordre de contiguïté
- Par exemple, la matrice de contiguïté à l'ordre  $k = 2$

$$W_2^* = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

# MATRICE DE CONTIGUÏTÉ D'ORDRE $k$

- ▶ Matrice à l'ordre  $k = 3$

MATRICE DE CONTIGUÏTÉ D'ORDRE  $k$ 

- ▶ Matrice à l'ordre  $k = 3$

$$W_3^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

- ▶ On peut alors construire la matrice de contiguïté **jusqu'à l'ordre 3**

MATRICE DE CONTIGUÏTÉ D'ORDRE  $k$ 

- Matrice à l'ordre  $k = 3$

$$W_3^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

- On peut alors construire la matrice de contiguïté **jusqu'à l'ordre 3**

$$W_3 = \begin{pmatrix} 0 & 1 & 2 & 1 & 1 & 3 \\ 1 & 0 & 2 & 1 & 1 & 3 \\ 2 & 2 & 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 & 1 & 3 \\ 1 & 1 & 1 & 1 & 0 & 2 \\ 3 & 3 & 1 & 3 & 2 & 0 \end{pmatrix}$$

## MATRICE DE DISTANCE À SEUIL

- ▶ Il s'agit de fixer une distance  $D$  au delà de laquelle l'interaction n'opère plus
  - ▶  $w_{ij} = 0$  si  $d_{ij} > D$
- ▶ En dessous de ce seuil, l'interaction existe
  - ▶  $w_{ij} = 1$  si  $d_{ij} < D$
- ▶ Par convention
  - ▶  $w_{ij} = 0$  si  $i = j$

⇒ Il s'agit donc d'une matrice basée sur la distance mais binaire et symétrique

# MATRICE DES PLUS PROCHES VOISINS

- ▶ Il s'agit de tenir compte des  $k$  plus proches voisins

- ▶  $w_{ij} = 1$  si  $i$  et  $j$  sont plus proche voisins

- ▶ S'ils ne sont pas plus proche voisins :

- ▶  $w_{ij} = 0$

- ▶ Par convention

- ▶  $w_{ij} = 0$  si  $i = j$

- ▶ On ne tient pas compte de la distance

- ▶ On regarde uniquement les  $k$  voisins de  $i$

⇒ Il s'agit donc d'une matrice binaire mais pas forcément symétrique

# MATRICE DE PONDÉRATION GÉNÉRALISÉE

- ▶ Les matrices binaires imposent une discontinuité dans l'interaction
  - ▶ e.g. les effets ne s'arrêtent pas toujours nettement à une distance donnée
- ▶ Les matrices généralisées ne sont pas binaires
  - ▶ elles dépendent généralement de la distance
  - ▶ elles peuvent aussi dépendre de paramètres estimés
  - ▶ elles dépendent parfois de distances non géographiques
    - ▶ e.g. fonction décroissante d'une distance économique ou sociale
    - ⇒ si  $r_i$  représente l'emploi d'une région  $i$

$$w_{ij} = (|r_i - r_j| + 1)^{-1}$$

# MATRICE DE PONDÉRATION GÉNÉRALISÉE

- ▶ Soit  $f(d_{ij})$  une fonction décroissante de la distance
  - ▶ Les poids de la matrices sont alors  $w_{ij} = f(d_{ij})$
- ▶ Les fonctions les plus fréquentes sont
  - ▶ matrice de distance inverse :  $f(d_{ij}) = d_{ij}^{-1}$
  - ▶ matrice de modèle de gravité :  $f(d_{ij}) = d_{ij}^{-2}$
  - ▶  $f(d_{ij}) = \exp(-2 \times d_{ij})$
- ▶ En généralisant on peut également considérer
  - ▶  $f(d_{ij}) = d_{ij}^{-\alpha}$
  - ⇒  $\alpha$  peut être estimé ou fixé

# MATRICE DE PONDÉRATION GÉNÉRALISÉE

- ▶ Il est possible de tenir compte de la taille des frontières  $b_{ij}$ 
  - ▶ les pondérations vont dépendre simultanément de  $d_{ij}$  et  $b_{ij}$
- ▶ Pondérations de Cliff et Ord (1973)
  - ▶  $w_{ij} = d_{ij}^{-\alpha} b_{ij}^{\beta}$
  - ▶ Ici,  $b_{ij}$  représente le pourcentage du périmètre de  $i$  constituant une frontière commune entre  $i$  et  $j$

# OPÉRATEUR DE DÉCALAGE SPATIAL

► Dans le domaine temporel, la notion de décalage est triviale :

- Une variable  $y_t$  est retardée d'une période :  $y_{t-1}$
- Une variable  $y_t$  est avancée d'une période :  $y_{t+1}$

⇒ Il n'existe que deux directions possibles

► Dans le domaine spatial, la transposition n'est pas parfaite

- Une variable  $x_i$  est décalée à gauche ou à droite : arbitraire
- Une variable  $x_i$  est décalée à l'est, à l'ouest, au nord, au sud : arbitraire

⇒ Il existe une infinité de direction pour un nombre indéterminé de voisins

⇒ La matrice de poids va jouer le rôle d'opérateur

# OPÉRATEUR DE DÉCALAGE SPATIAL

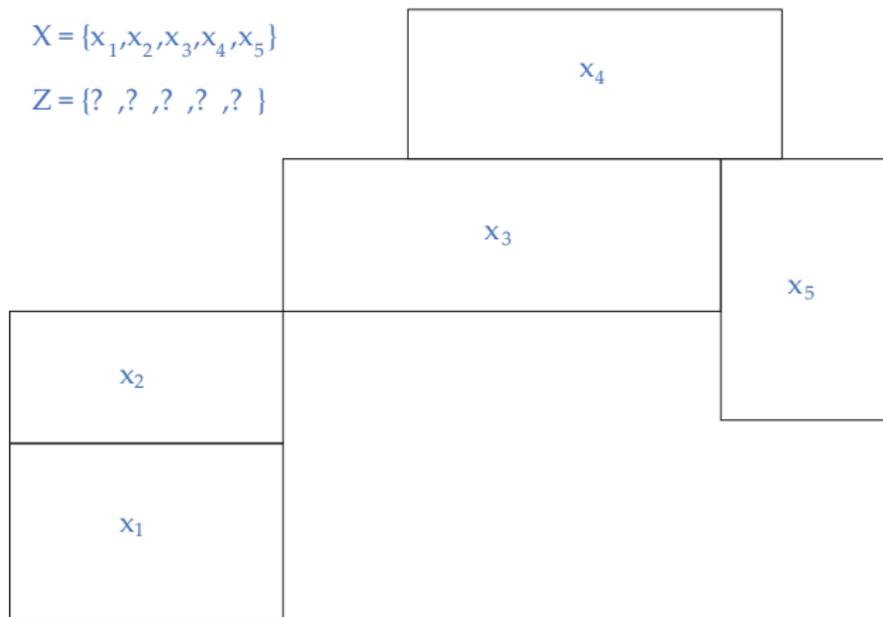
- ▶ Soit  $X = \{x_1, \dots, x_n\}$  et  $W$  une matrice de poids
  - ▶  $Z = WX$  est alors une variable spatialement décalée
  - ⇒ la matrice de poids agit comme un opérateur de décalage  $z_i = \sum_{j \neq i} w_{ij}x_j$
- ▶ Si la matrice de poids est standardisé,  $W^{(s)} = w_{ij} \left( \sum_j w_{ij} \right)^{-1}$ 
  - ⇒  $z_i$  représente la moyenne spatialement pondérée de  $x$  dans les  $j$  localités voisines de  $i$
  - ⇒ Attention, la standardisation en ligne détruit la symétrie
- ▶ Si on constate que  $z_i$  est similaire à  $(WX)_i$ 
  - ▶ association linéaire de valeurs élevées ou association linéaire de valeurs faibles
  - ⇒ autocorrélation spatiale positive
- ▶ Si on constate que  $z_i$  est dissimilaire à  $(WX)_i$ 
  - ▶ valeurs élevées chez  $z_i$  associées à valeurs faibles chez  $(WX)_i$
  - ⇒ autocorrélation spatiale négative

## EXEMPLE

- Sur la base d'une contiguïté de type **tour** à l'ordre 1 construire une variable spatialement décalée

$$X = \{x_1, x_2, x_3, x_4, x_5\}$$

$$Z = \{?, ?, ?, ?, ?\}$$



## EXEMPLE

- ▶ Contiguïté de type **tour** à l'ordre 1

## EXEMPLE

- Contiguïté de type **tour** à l'ordre 1

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad W^{(s)} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix}$$

- La variable spatialement décalée est alors  $Z = W^{(s)} X$

## EXEMPLE

- Contiguïté de type **tour** à l'ordre 1

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}, \quad W^{(s)} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix}$$

- La variable spatialement décalée est alors  $Z = W^{(s)} X$

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \\ 1/2x_4 + 1/2x_5 \\ 1/2x_3 + 1/2x_5 \\ 1/2x_3 + 1/2x_4 \end{pmatrix}$$