Chapitre 5 : L'autocorrélation spatiale

Nom de l'Auteur

2025-01-15

Introduction

- L'autocorrélation spatiale est un concept clé en statistique spatiale.
- Elle mesure dans quelle mesure les valeurs d'une variable, comme un taux de consultation moyen par commune, sont influencées par la localisation géographique.
- Objectifs :
  - Définir et comprendre l'autocorrélation spatiale.
  - Explorer l'indice de Moran, la matrice de poids (W), le diagramme de Moran et les LISA.

Comprendre l'autocorrélation spatiale

- **Définition** : L'autocorrélation spatiale décrit la corrélation d'une variable avec elle-même dans l'espace.
- **Pourquoi c'est important** :
  - Identifier des schémas géographiques significatifs (clusters).
  - Valider des hypothèses pour des modèles économiques ou environnementaux.
- **Types d'autocorrélation** :
  - Positive : Les valeurs similaires (élevées ou faibles) se regroupent.
  - Négative : Les valeurs dissemblables se regroupent.
  - Nulle : Absence de schéma spatial.

La matrice de poids (W)

Définition et rôle

- **Définition** : La matrice de poids formalise les relations spatiales entre les unités géographiques.
- **Rôle** : Elle sert de base pour calculer les indices spatiaux comme celui de Moran.

Types de matrice de poids

1. **Matrice binaire** :
   - $w_{ij} = 1$ si i et j sont voisins, sinon $w_{ij} = 0$.
   - Exemple : Deux communes sont voisines si elles partagent une frontière.
2. **Matrice pondérée par la distance** :
   - $w_{ij} = \frac{1}{d_{ij}^2}$, où d_{ij} est la distance entre i et j.
   - Une distance maximale (d_{max}) peut être définie pour limiter les connexions.
3. **Matrice pondérée par caractéristiques** :
   - Pondération selon des critères comme la proportion de frontières communes ou des similitudes démographiques.

Normalisation

- **Pourquoi normaliser ?** Pour uniformiser l'influence spatiale de chaque unité.
- **Formule** :

$$w_{ij}^{norm} = \frac{w_{ij}}{\sum_j w_{ij}}.$$

- **Effet** : Chaque ligne de la matrice a une somme égale à 1.

L'indice de Moran (I)

Formule générale

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \cdot \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Composantes de l'indice

- n : Nombre total d'unités spatiales.
- x_i, x_j : Valeurs observées pour les unités i et j.
- $\bar{x}$: Moyenne des observations.
- w_{ij} : Poids spatial entre i et j.

Interprétation

- $I > 0$: Autocorrélation positive (valeurs similaires regroupées).
- $I < 0$: Autocorrélation négative (valeurs opposées regroupées).
- $I = 0$: Absence d'autocorrélation.

Étapes de calcul

1. Déterminer la matrice W en fonction du voisinage spatial.
2. Calculer les écarts à la moyenne : $(x_i - \bar{x})$.
3. Multiplier ces écarts par w_{ij} pour toutes les paires (i, j).
4. Normaliser le résultat.

Diagramme de Moran

Définition

- Une visualisation graphique pour interpréter l'autocorrélation spatiale.

Construction

1. Axe x : Valeurs centrées $(x_i - \bar{x})$.
2. Axe y : Moyennes spatiales pondérées (Wy_i).
3. Quadrants :
    - **High-High** : Valeurs élevées entourées de valeurs élevées.
    - **Low-Low** : Valeurs faibles entourées de valeurs faibles.
    - **High-Low** : Valeurs élevées entourées de valeurs faibles.
    - **Low-High** : Valeurs faibles entourées de valeurs élevées.

Indicateurs locaux (LISA)

Définition

- Les LISA mesurent l'autocorrélation locale.
- Ils identifient des clusters ou des anomalies spécifiques.

Formule

$$I_i = \frac{x_i - \bar{x}}{\sigma^2} \cdot \sum_{j=1}^{n} w_{ij}(x_j - \bar{x})$$

Utilité des LISA

- Complètent l'analyse globale en identifiant des variations locales.
- Détectent des zones problématiques ou intéressantes pour des interventions ciblées.

Applications

- **Santé** : Identifier des clusters de consultations.
- **Urbanisme** : Repérer des zones de forte activité ou de déclin.
- **Environnement** : Analyser les poches de pollution ou de biodiversité.

Conclusion

- L'autocorrélation spatiale est un outil clé pour explorer les relations géographiques des données.
- La matrice de poids est essentielle pour définir les interactions spatiales.
- Les indices globaux et locaux, comme le Moran et les LISA, fournissent des perspectives complémentaires pour des analyses approfondies.