Exercice 4.3 : Démonstration de l'égalité entre l'Indice de Moran et la pente de régression du décalage spatial

Soit X un vecteur d'observations $x_1, x_2, \ldots, x_n$. Soit W une matrice de voisinage de dimension $n \times n$.

Hypothèses

La matrice W est normalisée par ligne : pour chaque ligne i, $\sum_{j=1}^{n} w_{ij} = 1$.

1. Égalité des moyennes ($\overline{WX} = \bar{x}$)

Pour démontrer que la pente de régression est égale à l'indice de Moran, nous devons d'abord prouver que la moyenne de la variable décalée (WX) est identique à celle de la variable d'origine (X).

Le décalage spatial au point i est défini par : $(WX)_i = \sum_{j=1}^{n} w_{ij} x_j$. La moyenne de ces décalages est :

$$\overline{WX} = \frac{1}{n} \sum_{i=1}^{n} (WX)_i = \frac{1}{n} \sum_{i=1}^{n} \left(\sum_{j=1}^{n} w_{ij} x_j \right)$$

En inversant l'ordre des sommations :

$$\overline{WX} = \frac{1}{n} \sum_{j=1}^{n} x_j \left(\sum_{i=1}^{n} w_{ij} \right)$$

Si l'on considère que la somme des poids par colonne est égale à 1 (cas d'une matrice bistochastique ou symétrique normalisée) :

$$\overline{WX} = \frac{1}{n} \sum_{j=1}^{n} x_j (1) = \bar{x}$$

2. Centrage du vecteur décalé (WX)

Montrons que si le vecteur X est centré, le vecteur spatialement décalé WX l'est également.

Supposons que la variable X est centrée. Par définition, cela signifie que sa moyenne est nulle :

$$\bar{x} = 0$$

D'après la propriété démontrée dans la section 1, nous savons que les moyennes sont égales ($\overline{WX} = \bar{x}$). En remplaçant $\bar{x}$ par 0, nous obtenons immédiatement :

$$\overline{WX} = 0$$

Conclusion : Si la matrice W est normalisée et si le vecteur X est centré, alors la moyenne du vecteur spatialement décalé est nulle. Le vecteur WX est donc également centré.

3. Expression de l'Indice de Moran (I)

Soit X un vecteur de n observations $x_1, x_2, \ldots, x_n$ associées à des unités spatiales.

La formule générale de l'indice de Moran est :

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\text{avec } S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}$$

Comme W est normalisée par ligne, alors $\frac{n}{S_0} = 1$. En sortant le terme $(x_i - \bar{x})$ de la somme interne sur j, on obtient :

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

4. Expression de I en fonction du décalage spatial (WX)

Le décalage spatial au point i est défini par : $(WX)_i = \sum_{j=1}^{n} w_{ij} x_j$.

Développons le terme entre crochets dans l'expression de I :

$$\sum_{j=1}^{n} w_{ij}(x_j - \bar{x}) = \sum_{j=1}^{n} w_{ij} x_j - \bar{x} \sum_{j=1}^{n} w_{ij}$$

Puisque $\sum_{j=1}^{n} w_{ij} = 1$, on a :

$$\sum_{j=1}^{n} w_{ij}(x_j - \bar{x}) = (WX)_i - \bar{x}$$

En remplaçant dans la formule de I, nous arrivons à :

$$I = \frac{\sum_{i=1}^{n}(x_i - \bar{x})((WX)_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{1}$$

5. Pente de la droite de régression (β)

Considérons la régression linéaire simple de WX sur X : $(WX)_i = \alpha + \beta x_i + \epsilon_i$.

La formule du coefficient de pente β par les Moindres Carrés Ordinaires (MCO) est :

$$\beta = \frac{Cov(X, WX)}{Var(X)} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})((WX)_i - \overline{WX})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Or, d'après la section 1, nous savons que $\overline{WX} = \bar{x}$. On en déduit :

$$\beta = \frac{\sum_{i=1}^{n}(x_i - \bar{x})((WX)_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2}$$

Conclusion

En comparant les équations (1) et (2), nous concluons que :

$$I = \beta$$

L'indice de Moran global correspond donc à la pente de la droite de régression linéaire de Moran.