

Chapitre 1

La matrice de pondération spatiale

Feuille d'exercices

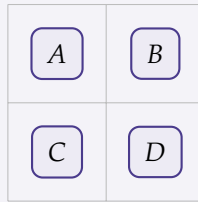
Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger
Exercices et corrigés · Économétrie spatiale (Master)

1 Exercices du chapitre

Exercice 1 — Contiguïté tour et contiguïté reine

On reprend les quatre régions en carré du cours : A et B côte à côte au nord, C et D côte à côte au sud, A au-dessus de C , B au-dessus de D .



- Écrire la matrice W (avant standardisation) sous critère de contiguïté **tour** (rook), en ordonnant les lignes et colonnes A, B, C, D .
- Écrire la matrice W sous critère de contiguïté **reine** (queen).
- Combien de voisins chaque région a-t-elle sous chacun des deux critères? Que devient, dans chaque cas, le poids w_{ij}^* après standardisation en ligne?

Exercice 2 — Calculer le décalage spatial Wy

On reprend la configuration en carré (contiguïté tour uniquement : $A \sim B, A \sim C, B \sim D, C \sim D$; pas de lien diagonal). Le revenu moyen par habitant (en milliers de dirhams) est $y_A = 12, y_B = 15, y_C = 9, y_D = 11$.

- Construire la matrice W^* standardisée en ligne sous ce critère.
- Calculer $(W^*y)_A, (W^*y)_B, (W^*y)_C$ et $(W^*y)_D$.
- Donner, en une phrase, l'interprétation économique de $(W^*y)_C$.
- La région C a le revenu le plus faible de l'échantillon ($y_C = 9$). Son décalage spatial $(W^*y)_C$ est-il, lui aussi, le plus faible des quatre? Commenter ce que cela révèle sur la configuration des voisinages.

Exercice 3 — Pondération par la distance

Les quatre régions occupent maintenant les coordonnées $A = (0, 1), B = (1, 1), C = (0, 0), D = (1, 0)$ (unité : centaines de kilomètres). Le taux de chômage régional est $y_B = 8, y_C = 10, y_D = 6$ (en %).

- Calculer les distances euclidiennes d_{AB}, d_{AC} et d_{AD} .
- Avec $\alpha = 1$ (poids $w_{ij} = 1/d_{ij}$), calculer les poids bruts w_{AB}, w_{AC}, w_{AD} , puis les poids standardisés en ligne $w_{AB}^*, w_{AC}^*, w_{AD}^*$.
- En déduire $(W^*y)_A$ et comparer au résultat obtenu dans l'exemple du cours avec la contiguïté reine ($(Wy)_A = 8$, poids égaux à $1/3$). Pourquoi les deux résultats diffèrent-ils, alors que les quatre mêmes régions sont en jeu?
- Que se passerait-il, qualitativement, sur le poids relatif de D (le plus éloigné de A) si l'on choisissait $\alpha = 2$ plutôt que $\alpha = 1$?

2 Exercice cumulatif (chapitre 1 et introduction)

Exercice 4 — De la matrice W au diagnostic de dépendance

L'introduction du cours distingue la dépendance substantielle (un mécanisme économique relie réellement les unités) de la corrélation nuisible (une variable omise crée l'apparence d'un lien), et précise que la conséquence sur les MCO diffère selon que la dépendance loge dans y ou dans le terme d'erreur. On reprend la configuration en carré et la contiguïté tour de l'exercice 2.

- (a) Un chercheur affirme : « la matrice W , à elle seule, permet de savoir si l'on est face à une dépendance substantielle ou à une corrélation nuisible ». Est-ce exact? Justifier à partir de ce que W formalise et de ce qu'elle ne formalise pas.
- (b) Pour la configuration en carré, un fonctionnaire régional propose de construire W à partir des flux commerciaux observés entre les quatre régions plutôt qu'à partir de la contiguïté géographique. D'après la mise en garde d'Anselin rappelée dans ce chapitre, pourquoi ce choix serait-il problématique s'il cherche ensuite à expliquer le PIB régional par ce même W ?
- (c) Le taux de chômage W_y calculé à l'exercice 2 (question b) sera l'ingrédient de droite du modèle SAR étudié au chapitre 3. Sans encore connaître l'équation exacte du modèle SAR, quelle est, intuitivement, la variable que ce modèle va chercher à expliquer à partir de W_y ?

3 Applications en sciences économiques

Exercice 5 — Choisir W pour des provinces marocaines

Une chercheuse étudie la corrélation spatiale du taux d'investissement direct étranger (IDE) entre les provinces du Maroc, dont certaines sont de très grande superficie (Sahara) et d'autres très petites et densément peuplées (Rif, Souss).

- (a) Pourquoi un critère de contiguïté simple (tour ou reine) risque-t-il d'être trompeur ici, compte tenu de l'hétérogénéité des superficies provinciales?
- (b) Pourquoi un critère de distance entre centroïdes présente-t-il le même risque, dans une moindre mesure?
- (c) Pourquoi le critère des k plus proches voisins est-il, dans ce contexte précis, souvent préféré par les praticiens? Rattacher la réponse au problème des « îles » mentionné dans le cours.

Exercice 6 — k plus proches voisins sur un exemple numérique

On reprend les quatre régions aux coordonnées de l'exercice 3 : $A = (0,1)$, $B = (1,1)$, $C = (0,0)$, $D = (1,0)$, et un indicateur de croissance du PIB régional $y_B = 3,2$, $y_C = 1,5$, $y_D = 2,0$ (en %).

- (a) Identifier les $k = 2$ plus proches voisins de A parmi $\{B, C, D\}$, à partir des distances calculées à l'exercice 3.
- (b) En déduire la matrice W_{2NN}^* standardisée en ligne pour la région A , puis calculer $(W_{2NN}^* y)_A$.
- (c) Le critère des k plus proches voisins n'est pas nécessairement symétrique. Donner, sans calcul supplémentaire, un exemple de configuration de points où j serait un plus proche voisin de i sans que i soit un plus proche voisin de j .