

Chapitre 9

Applications, régression géographiquement pondérée et limites

Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger
Polycopié de cours · Économétrie spatiale (Master)

Plan du chapitre

1. Un autre problème spatial : l'hétérogénéité
2. Le problème des unités spatiales modifiables
3. Les grands champs d'application
4. Ce qu'il faut retenir

1 Un autre problème spatial : l'hétérogénéité

1.1 Quand ce n'est pas la dépendance qui varie, mais l'effet lui-même

Un changement de perspective nécessaire

Tout ce cours, jusqu'à présent, traitait exclusivement la dépendance spatiale : les observations ne sont jamais réellement indépendantes entre elles. Un problème bien distinct, souvent confondu à tort avec le premier, est l'hétérogénéité spatiale : le coefficient β_k lui-même n'a peut-être pas la même valeur partout sur le territoire étudié. Ce chapitre distingue soigneusement les deux problèmes, avant de conclure sur les limites communes à ces deux approches.

1.2 La régression géographiquement pondérée

GWR, de Fotheringham, Brunsdon et Charlton

$$y_i = \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i$$

où (u_i, v_i) désignent les coordonnées géographiques propres de i . Un modèle local distinct est ainsi estimé en chaque point du territoire, par MCO pondérés au moyen d'un noyau qui accorde davantage de poids aux observations proches de (u_i, v_i) , et moins de poids aux observations plus lointaines.

Un rendement de l'éducation qui varie selon les régions

Le rendement salarial de l'éducation, déjà étudié au chapitre 10 de *Principes de microéconomie*, peut se révéler plus élevé dans une métropole, où l'offre d'emplois qualifiés demeure abondante, que dans une zone rurale où elle l'est nettement moins. La GWR estime directement cette variation, région par région, plutôt que de l'imposer identique partout comme le ferait nécessairement une régression classique.

Le choix délicat de la largeur de bande

Une bande étroite capture une variation locale fine, mais avec très peu d'observations disponibles par régression locale, donc une variance élevée. Une bande large lisse en revanche excessivement les résultats et se rapproche alors d'une régression globale unique, avec un biais certain si l'hétérogénéité est réellement présente. Le choix optimal s'obtient typiquement par validation croisée, en minimisant l'erreur de prédiction hors échantillon.

GWR, SAR et SEM répondent en réalité à deux questions différentes

Les modèles SAR et SEM supposent β constant partout, mais modélisent explicitement la dépendance existant entre observations. La GWR suppose au contraire l'indépendance complète des observations, mais laisse β varier librement dans l'espace considéré. Ces deux hypothèses peuvent d'ailleurs se révéler simultanément fausses — d'où des extensions récentes combinant les deux logiques à la fois, encore activement étudiées en recherche aujourd'hui.

2 Le problème des unités spatiales modifiables

2.1 Le résultat dépend toujours du découpage retenu

MAUP, formulé par Openshaw

Le problème des unités spatiales modifiables désigne le fait que les résultats d'une analyse spatiale dépendent directement du découpage administratif ou géographique retenu, lequel demeure le plus souvent arbitraire au regard du phénomène économique réellement étudié. Deux composantes distinctes le composent : l'effet d'échelle, où agréger en régions plus grandes change les corrélations mesurées, et l'effet de zonage, où à échelle égale, un découpage différent des mêmes unités change également les résultats obtenus.

Pourquoi ce problème n'admet aucune solution générale

Contrairement à un problème d'endogénéité ordinaire, pour lequel on peut toujours chercher un instrument valide, le MAUP n'admet aucune correction véritablement universelle : la seule parade disponible consiste à tester la robustesse des conclusions obtenues à plusieurs découpages plausibles distincts, et à rester prudent sur toute conclusion qui ne survivrait qu'à un seul découpage particulier — un écho direct de la mise en garde déjà formulée au chapitre 2 concernant le choix de W .

2.2 L'erreur écologique

Erreur écologique

Une corrélation observée entre variables agrégées au niveau régional ne se transpose jamais nécessairement au niveau individuel des personnes concernées. Une région affichant à la fois un fort taux d'immigration et un faible taux de chômage n'implique en rien que les immigrants de cette région précise soient moins souvent au chômage que les natifs eux-mêmes : la corrélation régionale observée peut fort bien masquer une composition différente des deux groupes considérés.

3 Les grands champs d'application

3.1 Où l'économétrie spatiale est aujourd'hui incontournable

Cinq domaines d'application aujourd'hui bien établis

Les modèles hédoniques de prix immobiliers, où la localisation constitue par construction même la variable explicative centrale de l'analyse. Les modèles de convergence régionale, qui testent si les régions les plus pauvres croissent effectivement plus vite — chapitre 8 de *Principes de macroéconomie* —, tout en tenant compte des spillovers de croissance observés entre régions voisines. La criminalité et sa diffusion spatiale bien réelle entre quartiers urbains. La concurrence fiscale stratégique entre collectivités locales. L'épidémiologie spatiale enfin, où les modèles SAR et SEM datent en réalité des mêmes années que leurs toutes premières applications en santé publique.

4 Ce qu'il faut retenir

Cinq points

La GWR traite l'hétérogénéité spatiale des coefficients eux-mêmes, un problème bien distinct de la dépendance spatiale déjà traitée par les modèles SAR et SEM. Le MAUP rappelle qu'aucun résultat spatial n'échappe jamais entièrement à l'arbitraire du découpage géographique retenu. L'erreur écologique rappelle de son côté qu'une corrélation agrégée ne dit strictement rien du niveau individuel des personnes concernées. Ces limites bien réelles ne disqualifient nullement l'économétrie spatiale dans son ensemble : elles imposent simplement la même discipline que partout ailleurs dans ce cours — tester systématiquement la robustesse, chiffrer précisément, et ne jamais confondre corrélation spatiale et véritable mécanisme causal sous-jacent. Le chapitre 10 referme enfin ce cours sur un bilan chiffré de l'ensemble du parcours suivi.

Pour aller plus loin

- [1] A. Stewart Fotheringham, Chris Brunsdon et Martin Charlton, *Geographically Weighted Regression*, Wiley, 2002.
- [2] Stan Openshaw, *The Modifiable Areal Unit Problem*, Geo Books, 1984.
- [3] William S. Robinson, « Ecological Correlations and the Behavior of Individuals », *American Sociological Review*, 1950.
- [4] Luc Anselin, *Spatial Econometrics : Methods and Models*, Kluwer Academic Publishers, 1988.
- [5] J. Paul Elhorst, *Spatial Econometrics : From Cross-Sectional Data to Spatial Panels*, Springer, 2014.