Plan

Modèle avec autocorrélation spatiale des erreurs

- Considérons un modèle en coupe transversale avec autocorrélation spatiale des erreurs

$$y_n = X_n\beta + \varepsilon_n \qquad \varepsilon_n = \lambda W_n \varepsilon_n + u_n$$

- Forme réduite :
  Si $I_n - \lambda W_n$ est inversible, la forme réduite du modèle SEM s'écrit comme suit :

$$y_n = X_n\beta + (I_n - \lambda W_n)^{-1} u_n$$

- La forme réduite implique qu'un choc aléatoire sur une observation affectera la valeur de la variable dépendante de l'ensemble des observations de l'échantillon grâce à la transformation spatiale inverse $(I_n - \lambda W_n)^{-1}$. C'est l'effet de diffusion d'un choc aléatoire. C'est une propriété des modèles SEM et SAR, mais le SEM n'inclut pas l'effet de multiplicateur spatial, passant par les variables explicatives.

Modèle SEM

- Ce modèle peut être interprété comme un modèle SDM contraint

$$(I_n - \lambda W_n)y_n = (I_n - \lambda W_n)(X_n\beta) + u$$

$$y = \lambda W_n y_n + X_n\beta - \lambda W X \beta + u$$

- qui est un modèle SDM

$$y = \lambda W_n y_n + X_n\beta + W_n X_n \gamma + \varepsilon_n$$

avec la contrainte non linéaire $\gamma = -\lambda\beta$.

Modèle SEM

- Filtre spatial
  Si λ était connu, on pourrait filtrer l'autocorrélation spatiale dans les X_n et y_n et utiliser un modèle classique pour estimer :
  $$(I_n - \lambda W_n)y_n = (I_n - \lambda W_n)X_n\beta + u$$
  λ généralement non connu et doit être estimé, comme β et σ_u^2.
- Matrice de variance-covariance
  en supposant que X_n est non stochastique, l'espérance de y_n est $\mathbf{E}(y_n) = X_n\beta$. Sa matrice de var-cov est
  $$V(y_n) = V(\varepsilon_n) = \sigma_u^2(I_n - \lambda W_n)^{-1}(I_n - \lambda W_n)'^{-1} = \sigma_u^2\Omega_\varepsilon(\lambda)$$
  C'est le même résultat que pour le modèle SAR. De nouveau, ses éléments diagonaux ne sont pas constants et les éléments hors-diagonaux sont non nuls.

Plan

Impacts dans le modèle SEM

- Le modèle SEM s'écrit :

$$y_n = \alpha\iota_n + X_n\beta + \varepsilon_n \qquad \varepsilon_n = \lambda W_n\varepsilon_n + u_n$$

qui peut s'écrire comme

$$y_n = \alpha\iota_n + X_n\beta + (I_n - \lambda W_n)^{-1}u_n$$

Les dérivées partielles de y_n par rapport à X_k pour $k = 1, \cdots, K$:

$$\frac{\partial y}{\partial X_k'} = S_k(W) = I_N\beta_k$$

donnent les mêmes résultats que dans un modèle linéaire classique standard. Il n'y a pas de spillovers spatiaux.

Hypothèses des MCO

- Soit le modèle de régression $y = X\beta + \varepsilon$
- 3 corps d'hypothèses
  1. Sur les variables explicatives : non-stochastiques, de rang complet
  2. Sur les erreurs : $\mathbf{E}(\varepsilon) = 0, \quad \mathbf{E}(\varepsilon\varepsilon') = \sigma^2 I_n$
  3. Modèle linéaire (y est une combinaison linéaire des variables indépendantes)

Estimation par les MCO

- Minimisation de la somme des résidus (RSS) par rapport à β:

$$RSS = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon = (y - X\beta)'(y - X\beta)$$

- On obtient

$$\hat{\beta}_{MCO} = (X'X)^{-1}X'y \quad (9)$$

$$\hat{\sigma}^2 = \frac{e'e}{n-k} \quad (10)$$

- Propriétés statistiques de β
  - Estimateur centré : $\mathbf{E}(\hat{\beta}) = \beta_0$
  - Estimateur BLUE
  - Estimateur convergent

MCG purs

- Soit le modèle de régression $y = X\beta + \varepsilon$
- La matrice de variance-covariance des erreurs non sphérique: $\mathbf{E}(\varepsilon\varepsilon') = \sigma^2 V$ (V définie positive et connue)
- Application des MCO donne un estimateur centré, convergent mais inefficient.
- Estimateur des MCG

$$\hat{\beta}_{MCG} = (X'V^{-1}X)^{-1}X'V^{-1}y$$

- Estimateur équivalent à l'estimateur des MCO sur le modèle transformé

$$Py = PX\beta + P\varepsilon$$

où P est une matrice non singulière tq $PVP' = I \Rightarrow V^{-1} = P'P$

MCG réalisables

- V n'est pas connu en pratique. Il faut l'estimer
- Supposons que V dépend d'un vecteur de paramètres inconnus θ_0 de dimension $m \times 1$.
- Estimation en 2 étapes
  1. Estimation convergence du vecteur des paramètres θ_0 dont dépend V
  2. Application des MCG avec $\hat{V}$ à la place de V

$$\hat{\beta}_{MCGR} = (X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y$$

  3. Sous certaines conditions assez générales, on peut montrer que l'estimateur des MCG réalisables est asymptotiquement équivalent à l'estimateur des MCG purs.

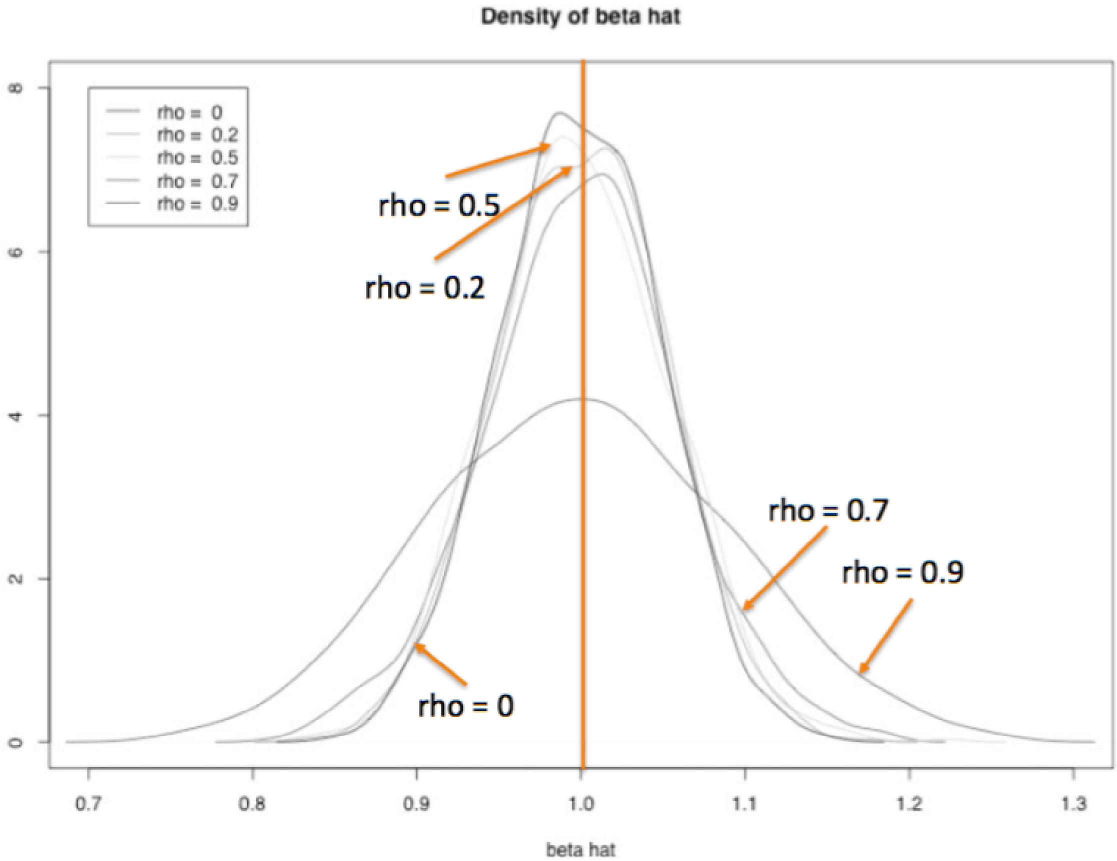
Estimation du modèle SEM

$$y_n = X_n\beta + \varepsilon_n \qquad \varepsilon_n = \lambda W_n \varepsilon_n + u_n \quad u_n \sim N(0, \sigma_u^2 I)$$
$$y_n = X_n\beta + \varepsilon_n \quad \varepsilon \sim N(0, \sigma_u^2 \Omega_\varepsilon(\lambda))$$
$$\sigma_u^2 \Omega_\varepsilon(\lambda) = \sigma_u^2 (I_n - \lambda W_n)^{-1} (I_n - \lambda W_n)'^{-1}$$

- Erreurs non sphériques ($\sigma_u^2 \Omega_\varepsilon(\lambda) \neq \sigma_u^2 I$)
- **Estimateur MCO reste centré mais plus efficace**
- **Inférence statistique fondée sur l'estimateur MCO est biaisée**
- Solution théorique : MCG réalisables
  Problème: l'estimateur des λ par les MCO n'est pas convergent pour le modèle autorégressif spatial
- Nécessité d'appliquer une autre méthode d'estimation pour aboutir à des estimateurs convergents, asymptotiquement efficients et normaux.

$\Rightarrow$ Estimation par MV (ou GMM)

Estimateur des MCO de β



Estimation du SEM par Max de Vraisemblance I

Fonction de vraisemblance pour une loi normale multivariée avec $\varepsilon \sim N(0, \Sigma)$:

$$\ln L = -\frac{n}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma| - \frac{1}{2}\varepsilon'\Sigma^{-1}\varepsilon$$

- Pour le SEM, $\Sigma = \sigma_u^2\Omega_\varepsilon(\lambda)$ et $|\sigma_u^2\Omega_\varepsilon(\lambda)| = \left(\sigma_u^2\right)^N |I - \lambda W|^{-2}$, La fonction de log-vraisemblance s'écrit

$$\begin{aligned}
\ln L(\beta', \lambda, \sigma_\epsilon^2) &= -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\sigma_u^2\Omega_\varepsilon(\lambda)| \\
&\quad - \frac{1}{2\sigma_u^2}(y - X\beta)'\Omega_\varepsilon(\lambda)^{-1}(y - X\beta) \\
\ln L(\beta', \lambda, \sigma_\epsilon^2) &= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma_u^2) + \ln|I - \lambda W| \\
&\quad - \frac{1}{2\sigma_u^2}(y - X\beta)'\Omega_\varepsilon(\lambda)^{-1}(y - X\beta)
\end{aligned}$$

Estimation du SEM par Max de Vraisemblance II

- Nécessaire d'imposer l'hypothèse 4 pour garantir l'existence du Jacobien de la transformation.

$\Rightarrow$ Tout comme dans le modèle SAR, normalisation de la matrice d'interactions pour obtenir des espaces de paramètres comparables entre modèles.

Fonction de log-vraisemblance

- En notant que $\Omega_\varepsilon(\lambda)^{-1} = (I - \lambda W)'(I - \lambda W)$

$$
\begin{aligned}
\ln L(\beta', \lambda, \sigma_\epsilon^2) &= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma_u^2) + \ln|I - \lambda W| \\
&\quad - \frac{1}{2\sigma_u^2}\left[(I - \lambda W)(y - X\beta)\right]'\left[(I - \lambda W)(y - X\beta)\right] \\
\ln L(\beta', \lambda, \sigma_\epsilon^2) &= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma_u^2) + \ln|I - \lambda W| \\
&\quad - \frac{1}{2\sigma_u^2}\left[(I - \lambda W)y - (I - \lambda W)X\beta\right]'\left[(I - \lambda W)y - (I - \lambda W)X\beta\right] \\
\ln L(\beta', \lambda, \sigma_\epsilon^2) &= -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma_u^2) + \ln|I - \lambda W| \\
&\quad - \frac{1}{2\sigma_u^2}\left[y_L - X_L\beta\right]'\left[y_L - X_L\beta\right]
\end{aligned}
$$

où $X_L = (I_n - \lambda W_n)X_n$ and $y_L = (I_n - \lambda W_n)y_n$.

Estimation par Max de vraisemblance

- Supposons que λ est connu, les CPO pour β et σ_u^2 sont :

$$\frac{\partial \ln L}{\partial \beta'} = -\frac{1}{2\sigma_u^2}(2X_L'y_L + 2\beta X_L'X_L)$$
$$\frac{\partial \ln L}{\partial \sigma_u^2} = -\frac{N}{\sigma_u^2} + \frac{1}{\sigma_u^4}(y_L - X_L\beta)'(y_L - X_L\beta)$$

- L'estimateur du maximum de vraisemblance de β et σ_u^2, étant donné λ, sont obtenus comme:

$$\hat{\beta}_{MV}(\lambda) = (X_L'X_L)^{-1}X_L'y_L \quad (11)$$
$$\hat{\sigma}^2_{MV}(\lambda) = \frac{1}{N}(y_L - X_L\beta)'(y_L - X_L\beta) \quad (12)$$

Estimation par Max de vraisemblance

- En substituant (11) et (12) dans la fonction de log-vraisemblance, on obtient une fonction de log-vraisemblance qui ne dépend non linéairement que d'un seul paramètre λ :

$$\ln L = -\frac{n}{2}[1 + \ln(2\pi)] - \frac{n}{2}\ln\left(\frac{\hat{u}'\hat{u}}{N}\right) + \sum_{i=1}^{n}\ln(1 - \lambda\omega_i) \qquad (13)$$

  où $\hat{u} = y_L - X_L\hat{\beta}_{MV}(\lambda)$. Donc, $\hat{u}'\hat{u} = y_L'Y_L - y_L'X_L(X_L'X_L)^{-1}X_L'y_L$, où $y_L = y - \lambda Wy$ et $X_L = X - \lambda WX$ sont les variables spatialement filtrées.
- L'estimateur du max de vraisemblance de λ est obtenu par optimisation numérique de la fonction de log-vraisemblance concentrée.

Estimation par Maximum de vraisemblance I

- La procédure d'estimation est plus complexe que pour le SAR car $\hat{u}$ dans la log-vraisemblance concentrée est une fonction indirecte de λ car $\hat{\beta}$ est obtenu pour une valeur donnée de λ. Ainsi, une seule optimisation numérique de la fonction de log-vraisemblance concentrée pour λ ne suffit pas pour obtenir les valeurs estimées de tous les autres paramètres inconnus.
- Une approche itérative est nécessaire. On alterne entre une estimation de λ conditionnelle à un vecteur de résidus $\hat{u}$ obtenus pour une valeur de β et une estimation de β et σ^2 conditionnelles à une valeur de λ, jusqu'à obtenir une convergence numérique.

Estimation par Maximum de vraisemblance

La procédure d'estimation est comme suit

1. Régresser y sur X par les MCO et calculer les résidus $\hat{u} = y - X\beta_{OLS}$.

Estimation par Maximum de vraisemblance

La procédure d'estimation est comme suit

1. Régresser y sur X par les MCO et calculer les résidus $\hat{u} = y - X\beta_{OLS}$.
2. Etant donné $\hat{u}$, trouver λ qui maximise la fonction de log-vraisemblance.

Estimation par Maximum de vraisemblance

La procédure d'estimation est comme suit

1. Régresser y sur X par les MCO et calculer les résidus $\hat{u} = y - X\beta_{OLS}$.
2. Etant donné $\hat{u}$, trouver λ qui maximise la fonction de log-vraisemblance.
3. Etant donné λ, effectuer une régression MCGR pour obtenir $\hat{\beta}_{MCGR}$.

Estimation par Maximum de vraisemblance

La procédure d'estimation est comme suit

1. Régresser y sur X par les MCO et calculer les résidus $\hat{u} = y - X\beta_{OLS}$.
2. Etant donné $\hat{u}$, trouver λ qui maximise la fonction de log-vraisemblance.
3. Etant donné λ, effectuer une régression MCGR pour obtenir $\hat{\beta}_{MCGR}$.
4. Calculer ensuite un nouveau vecteur de résidus $\hat{u} = y - X\hat{\beta}_{MCGR}$.

Estimation par Maximum de vraisemblance

La procédure d'estimation est comme suit

1. Régresser y sur X par les MCO et calculer les résidus $\hat{u} = y - X\beta_{OLS}$.
2. Etant donné $\hat{u}$, trouver λ qui maximise la fonction de log-vraisemblance.
3. Etant donné λ, effectuer une régression MCGR pour obtenir $\hat{\beta}_{MCGR}$.
4. Calculer ensuite un nouveau vecteur de résidus $\hat{u} = y - X\hat{\beta}_{MCGR}$.
5. S'il y a convergence numérique, càd que les valeurs des résidus et de $\hat{\beta}_{MCGR}$ ne changent plus d'itérations en itérations, alors étant donné $\hat{u}$ et $\hat{\lambda}_{ML}$, on calcule $\hat{\sigma}^2_{ML}$,

Estimation par Maximum de vraisemblance

La procédure d'estimation est comme suit

1. Régresser y sur X par les MCO et calculer les résidus $\hat{u} = y - X\beta_{OLS}$.
2. Etant donné $\hat{u}$, trouver λ qui maximise la fonction de log-vraisemblance.
3. Etant donné λ, effectuer une régression MCGR pour obtenir $\hat{\beta}_{MCGR}$.
4. Calculer ensuite un nouveau vecteur de résidus $\hat{u} = y - X\hat{\beta}_{MCGR}$.
5. S'il y a convergence numérique, càd que les valeurs des résidus et de $\hat{\beta}_{MCGR}$ ne changent plus d'itérations en itérations, alors étant donné $\hat{u}$ et $\hat{\lambda}_{ML}$, on calcule $\hat{\sigma}^2_{ML}$,
6. Sinon, remaximiser la fonction de log-vraisemblance concentrée pour obtenir un nouveau λ.

Estimation par Max de Vraisemblance

- Tout comme pour le modèle SAR, sous les hypothèses de régularité de Lee (2004), l'estimateur du max de vraisemblance possède les bonnes propriétés asymptotiques : convergence, normalité et efficience asymptotique
- En notant $G_n = W(I_n - \lambda_0 W_n)^{-1}$, on écrit la matrice de Var-cov comme : AsyVar$[\beta', \lambda, \sigma^2] =$

$$\begin{bmatrix} \frac{1}{\sigma^2} X_L' X_L & 0 & 0 \\ 0 & \mathrm{tr}(G_n^2) + \mathrm{tr}(G_n' G_n) & \frac{1}{\sigma^2} tr(G_n) \\ 0 & \frac{1}{\sigma^2} tr(G_n) & \frac{n}{2\sigma^4} \end{bmatrix}^{-1} \tag{14}$$

- Cette matrice est bloc-diagonale. Ainsi, la précision de λ n'affecte pas la précision de l'estimateur des autres paramètres.

Plan

Modèle SLX

- Ce modèle s'écrit

$$y_n = \alpha \iota_n + \sum_{k=1}^{K} \beta_k X_k + \sum_{k=1}^{K} \gamma_k W_n X_k + \varepsilon_n = \sum_{k=1}^{K} (I_n \beta_k + W_n \gamma_k) X_k + \varepsilon_n$$

- Modèle avec spillovers locaux. Ce sont uniquement les variables explicatives de l'observation *i* et de ses voisins qui affectent la valeur de la variable dépendante de *i*.
- Ce modèle peut être estimé par les MCO

Analyse d'impacts

Les matrices d'impacts de ce modèle s'écrivent (pour k= 1, ..., K)

$$\frac{\partial y_n}{\partial X'_k} = S_k(W_n) = I_n\beta_k + W_n\gamma_k$$

Les éléments diagonaux de cette matrice représentent les effets directs qui sont constants (β_k) alors que les éléments hors diagonaux représentent les effets indirects. Cependant, contrairement au modèle SAR, ces effets indirects sont directement issus de la matrice W_n.