誤差が定常性をもつ回帰モデルの最尤推定量は漸近有効か

統計数理研究所　力丸佑紀

データサイエンスコンソーシアム　柴田里程

空間データモデル

- 空間ラグモデル (Spatial Lag Model, SLM)

$$\boldsymbol{y} = \lambda W \boldsymbol{y} + X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

- 空間誤差モデル (Spatial Error Model, SEM)

$$\begin{cases} \boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u} \\ \boldsymbol{u} = \lambda W \boldsymbol{u} + \boldsymbol{\varepsilon} \end{cases}$$

- パラメータ $\boldsymbol{\theta}^{\top} = (\boldsymbol{\beta}^{\top}, \lambda, \sigma)$
  - 回帰項 $X\boldsymbol{\beta}$
  - 空間重み行列 W，空間自己回帰パラメータ λ
  - $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 I)$

空間ラグモデルの書き換え

- 空間ラグモデル (Spatial Lag Model, SLM)

$$\boldsymbol{y} = \lambda W\boldsymbol{y} + X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\Leftrightarrow \quad (I - \lambda W)\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\Leftrightarrow \quad \boldsymbol{y} = (I - \lambda W)^{-1} X\boldsymbol{\beta} + (I - \lambda W)^{-1}\boldsymbol{\varepsilon}$$

$$\Leftrightarrow \quad \begin{cases} \boldsymbol{y} = X(\lambda)\boldsymbol{\beta} + \boldsymbol{u} \\ \boldsymbol{u} = \lambda W\boldsymbol{u} + \boldsymbol{\varepsilon} \end{cases}$$

$X(\lambda) = (I - \lambda W)^{-1}X$, $\boldsymbol{u} = (I - \lambda W)^{-1}\boldsymbol{\varepsilon}$ とおく

- SLM は SEM の変形（X が $X(\lambda)$ に置き換わった）

SLMのパラメータ推定

- Lee(2004) 最尤法によるパラメータ推定

$$L = -\log\det(I - \lambda W) - \frac{N}{2}\log 2\pi\sigma^2 + \frac{1}{2\sigma^2}(\boldsymbol{y} - X(\lambda)\boldsymbol{\beta})^\top (I - \lambda W)^\top (I - \lambda W)(\boldsymbol{y} - X(\lambda)\boldsymbol{\beta})$$

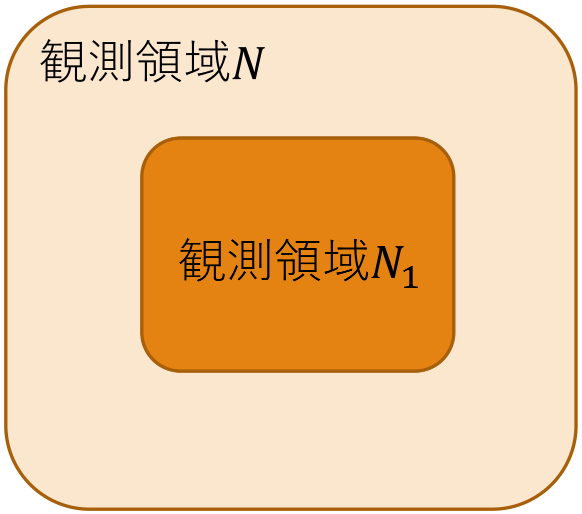
- 「漸近有効推定量が得られる」との結果

ホント？？

誤差 $\boldsymbol{u}$ の一貫性

- Lee(2004)

$$E(\boldsymbol{u}\boldsymbol{u}^{\top}) = \sigma^2(I-\lambda W)^{-1}\{(I-\lambda W)^{-1}\}^{\top}$$



- 観測領域 N での $\boldsymbol{u}$ の分散共分散行列

$\sigma^2(I-\lambda W)^{-1}\{(I-\lambda W)^{-1}\}^{\top}$　の一部？

- 観測領域 N_1 での $\boldsymbol{u}_1$ の分散共分散行列

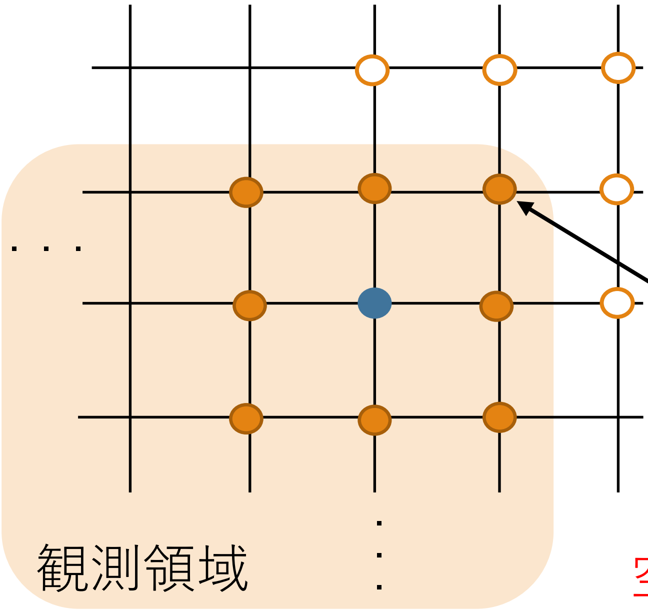
$\sigma^2(I-\lambda W_1)^{-1}\{(I-\lambda W_1)^{-1}\}^{\top}$

→ 観測領域が変わるたびに $\boldsymbol{u}$ の分布が変わる

$\boldsymbol{u}$ の一貫性のなさの原因

◆ エッジエフェクト

たとえば,
空間重み行列 W を隣接行列に設定
「青●は橙●の影響を受ける」

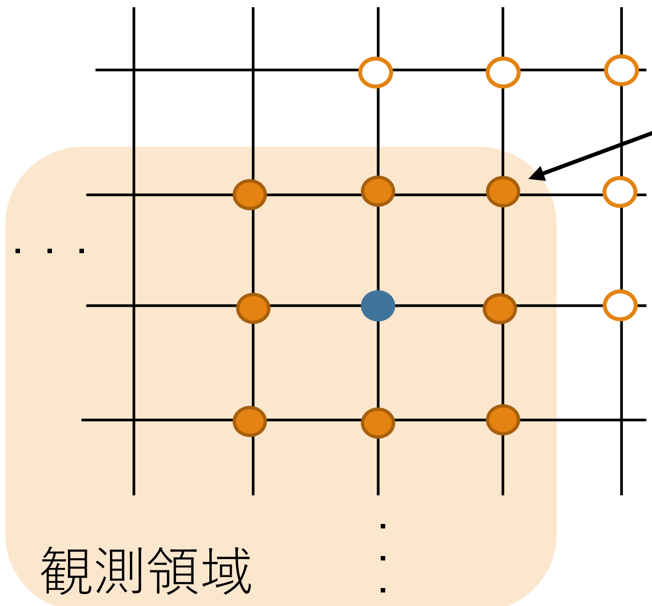


行列で表現された SLM や SEM では
観測領域の端ではそうならない

「この値は橙〇から影響を受けない」

空間データだからエッジエフェクトが大きい
（実際，数式でも大きな問題が起きる）

$\boldsymbol{u}$ に一貫性をもたせる



実際のデータ生成では
「この橙●も橙○の影響を受ける」
でないと不自然.

- 一貫性をもたせる方法のひとつ

  観測領域を変えても,
  観測領域の端を見ても,
  観測領域の外を見ても,
  同じ構造になっていれば…

◆ 誤差 $\boldsymbol{u}$ は弱定常性をもつ SAR モデルに従う

$$\underline{P(T_1, T_2)}u_v = \varepsilon_v,\ \ \varepsilon_v \sim N(0, \sigma^2) \Longrightarrow E(u_v u_{v'}) = \gamma_{v-v'}$$

$I - \lambda W$ と対応

弱定常性 $\gamma_{v,v'} = \gamma_{v-v'}$

- データ生成メカニズム

$$\begin{cases} \boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u} \\ P(T_1, T_2)u_v = \varepsilon_v \\ \varepsilon_v \sim N(0, \sigma^2) \end{cases}$$

$$P(T_1, T_2) = \lambda \sum_{k \in K} w_k T_1^{k_1} T_2^{k_2}$$

$$E(u_v u_{v'}) = \gamma_{v-v'}$$

- モデル

$$\begin{cases} \boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{u} \\ (I - \lambda W)\boldsymbol{u} = \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} \sim N(0, \sigma^2 I) \end{cases}$$

  - 尤度

$$L = -\log\det(I - \lambda W) - \frac{N}{2}\log 2\pi\sigma^2 + \frac{1}{2\sigma^2}(\boldsymbol{y} - X(\lambda)\boldsymbol{\beta})^\top (I - \lambda W)^\top (I - \lambda W)(\boldsymbol{y} - X(\lambda)\boldsymbol{\beta})$$

誤差が定常性をもつ回帰モデルの最尤推定量は漸近有効か

漸近バイアス

- λ と σ では漸近バイアスあり
- $\boldsymbol{\beta}$ では漸近バイアスなし

$$E\left(\frac{1}{\sqrt{N}}\frac{\partial L}{\partial \lambda}\right) = \frac{1}{\sqrt{N}}\mathrm{tr}\big((I-\lambda W)W\big) - \frac{1}{\sqrt{N}\sigma^2}\underline{\mathrm{tr}\big((I-\lambda W)^\top W \quad \Sigma \quad\big)}$$

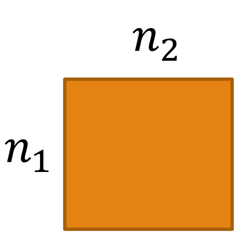
要素 $\gamma_{v,v'} = \gamma_{v-v'}$

前の項と打ち消し合う

$1/\sqrt{N}$ 倍すれば消える

$$(n_1 - |k_1 - k_1'|)(n_2 - |k_2 - k_2'|)\gamma_{-k_1+k_1',-k_2+k_2'} + \cdots$$

$O\left(\sqrt{\frac{n_2}{n_1}} + \sqrt{\frac{n_1}{n_2}}\right)$ で残る



$\boldsymbol{\beta}$ の漸近分散

- $\mathrm{Cov}\left(\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right) = J^{-1} E(\boldsymbol{g}\boldsymbol{g}^T)(J^{-1})^\top$

$$\boldsymbol{g} = \frac{1}{\sqrt{N}} \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \quad J = \frac{1}{N} \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$$

  - $E(\boldsymbol{g}\boldsymbol{g}^T) = \begin{pmatrix} A & \boldsymbol{b} & \boldsymbol{c} \\ \boldsymbol{b}^\top & d & e \\ \boldsymbol{c}^\top & e & f \end{pmatrix}$

漸近バイアスの影響で $O\left(\sqrt{\frac{n_2}{n_1}} + \sqrt{\frac{n_1}{n_2}}\right)$

- $\boldsymbol{\beta}$ も含む推定量の漸近分散は

大きくなる，場合によっては発散

漸近有効性を満たす近似尤度

$$L_A = \frac{1}{2}\log\det(A) - \frac{N}{2}\log 2\pi - \frac{1}{2}\left(\boldsymbol{y} - X(\lambda)\boldsymbol{\beta}\right)^{\top} \boxed{\tilde{A}}\left(\boldsymbol{y} - X(\lambda)\boldsymbol{\beta}\right)$$

$$A = \frac{\lambda^2}{\sigma^2} \sum_{\boldsymbol{k},\boldsymbol{k}' \in \mathcal{K}} w_{\boldsymbol{k}} w_{\boldsymbol{k}'} C_{n_1}^{-k_1+k_1'} \otimes C_{n_2}^{-k_2+k_2'}$$

$$\tilde{A} = \frac{\lambda^2}{\sigma^2} \sum_{\boldsymbol{k},\boldsymbol{k}' \in \mathcal{K}} w_{\boldsymbol{k}} w_{\boldsymbol{k}'} \boxed{d_{n_1}^{|k_1-k_1'|} d_{n_2}^{|k_2-k_2'|}} C_{n_1}^{-k_1+k_1'} \otimes C_{n_2}^{-k_2+k_2'}$$

縮小率 $d_n = 1 + \frac{1}{n}$

- エッジエフェクトを縮小率で無視できるように調整すればOK

まとめ

- SLM, SEM の最尤推定量は漸近有効だと信じられてきた（Lee(2004)）.
  - しかし，それは誤差 $\boldsymbol{u}$ に一貫性のない分布を想定したものだった.
- 一貫性をもたせるために $\boldsymbol{u}$ に弱定常性を仮定（現実に則した仮定）
  - すると，λ，σ に漸近バイアスあり
  - 漸近バイアスが $\boldsymbol{\beta}$ の漸近分散にもきいてくる
- パラメータ推定には L_A を.