

空間ラグモデルの問題 とその解決策

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, 慶應義塾大学 柴田里程

空間ラグモデル

- Spatial Lag Model, SLM

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

自己回帰項 回帰項

W : 空間重み行列, $\boldsymbol{\varepsilon} \sim N(0, \sigma^2 I)$

- Reduced form

$$\mathbf{y} = X(\lambda) \boldsymbol{\beta} + \mathbf{u}, \quad S(\lambda) \mathbf{u} = \boldsymbol{\varepsilon}$$

回帰モデルの誤差

$$S(\lambda) = I - \lambda W, \quad X(\lambda) = S(\lambda)^{-1} X$$

これまでのSLMのパラメータ推定

- 対数尤度(Lee,2004)

$$L_L(\boldsymbol{\beta}, \lambda, \sigma^2) = -\frac{1}{2} \log \det \left(\frac{S(\lambda)^\top S(\lambda)}{\sigma^2} \right) - \frac{N}{2} \log 2\pi \\ + \frac{1}{2\sigma^2} \mathbf{u}^\top S(\lambda)^\top S(\lambda) \mathbf{u}$$

- 漸近有効推定量を得られる



SLMの問題

- 誤差 $\mathbf{u}$ の非一貫性

$S(\lambda)\mathbf{u} = \boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon} \sim N(0, \sigma^2 I)$ より

$$\begin{aligned} E(\mathbf{u}\mathbf{u}^T) &= \sigma^2 S(\lambda)^{-1} \{S(\lambda)^{-1}\} \\ &= \sigma^2 (I - \lambda W)^{-1} \{(I - \lambda W)^{-1}\}^T \end{aligned}$$



- 観測領域 N での $\mathbf{u}$ の分散共分散行列

$$\sigma^2 (I - \lambda W)^{-1} \{(I - \lambda W)^{-1}\}^T$$

- 観測領域 N_1 での $\mathbf{u}_1$ の分散共分散行列

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

→ 観測領域が変わるたびに $\mathbf{u}$ の分布が変わる

誤差 $\mathbf{u}$ に定常性を仮定

- $\mathbf{u}$ の一貫性を確保する方法のひとつ

➤ 誤差 $\mathbf{u}$ に定常性を仮定 $E(u_v u_{v'}) = \gamma_{v-v'}$

データ生成メカニズム

モデル (SLM)

$$\begin{cases} \mathbf{y} = X(\lambda)\boldsymbol{\beta} + \mathbf{u} \\ \underline{P(T_1, T_2)u_v = \varepsilon_v} \end{cases}$$

$$\begin{cases} \mathbf{y} = X(\lambda)\boldsymbol{\beta} + \mathbf{u} \\ S(\lambda)\mathbf{u} = \boldsymbol{\varepsilon} \end{cases}$$

$$1 - \lambda \sum_{k \in K} w_k T_1^{k_1} T_2^{k_2}$$

対数尤度 L_L から得られる推定量は

最悪, 漸近バイアスが発散する.

漸近バイアス

- W が隣接行列の場合

➤ $W = (U_{n_1} + U_{n_1}^{-1}) \otimes I_{n_2} + I_{n_1} \otimes (U_{n_2} + U_{n_2}^{-1})$

n_2

観測領域

- U_n は $(j, j+1)$ 要素のみ 1, 他は 0 の行列

n_1

$$\sqrt{NE}(\hat{\lambda} - \lambda)$$

$$\approx \left(\sqrt{\frac{n_2}{n_1}} + \sqrt{\frac{n_1}{n_2}} \right) \frac{-\sum_{t \neq 0} \gamma_{1,2t} - \sum_{t \neq 0, -1} \gamma_{0,2t+1} - \lambda \left(3 \sum_{t \neq 0, 1} \gamma_{0,2t} + \sum_{t \neq 0} \gamma_{2,2t} + 8 \sum_{t \neq 0, -1} \gamma_{1,2t+1} \right)}{\frac{1}{N} \boldsymbol{\beta}^\top X(\lambda)^\top W^\top W X(\lambda) \boldsymbol{\beta}}$$

- $\hat{\boldsymbol{\beta}}$ の漸近バイアスにも影響

$$\sqrt{NE}(\hat{\boldsymbol{\beta}}(\hat{\lambda}) - \boldsymbol{\beta}) \approx E \left(\frac{\partial \hat{\boldsymbol{\beta}}(\lambda)}{\partial \lambda} \right) \cdot \sqrt{NE}(\hat{\lambda} - \lambda)$$

解決策

- 近似尤度 L_A を使う.

$$L_A = \frac{1}{2} \log \det(A) - \frac{N}{2} \log 2\pi - \frac{1}{2} \mathbf{u}^\top \tilde{A} \mathbf{u}$$

$$\text{ただし, } A = \frac{1}{\sigma^2} (I - \lambda \tilde{W})^T (I - \lambda \tilde{W}), \quad \tilde{W} = \sum_{k \in K} w_k \underbrace{C_{n_1}^{k_1} \otimes C_{n_2}^{k_2}}_{\text{巡回行列}},$$

$$\tilde{A} = \frac{1}{\sigma^2} \left(I - \lambda \sum_{k \in K} w_k \underbrace{d_{n_1}^{|k_1 - k'_1|} d_{n_2}^{|k_2 - k'_2|}}_{d_n = 1 + 1/n} \left(C_{n_1}^{k_1 - k'_1} \otimes C_{n_2}^{k_2 - k'_2} + C_{n_1}^{-k_1 + k'_1} \otimes C_{n_2}^{-k_2 + k'_2} \right) \right)$$

- 漸近バイアスなし
- 漸近有効性ももつ

見過ごされてきた問題

- 最尤推定量の一致性は前提とした.
 - u の一貫性とは無関係.
 - L_L についても, L_A についても.
- 実は, そう簡単に証明できることではない.

一致性の証明

- Ferguson(1996) (i.i.d. ver)

$$U_N(\theta_0, \theta) = \frac{1}{N} \log \frac{f(y, \theta)}{f(y, \theta_0)}, \quad \mu(\theta_0, \theta) = \frac{1}{N} E_{\theta_0} \frac{f(y, \theta)}{f(y, \theta_0)}$$

$$U_N(\theta_0, \theta) \rightarrow \mu(\theta_0, \theta) \quad a.s. \text{ as } N \rightarrow \infty$$

ならば

$$\hat{\theta} \rightarrow \theta_0 \quad \text{in prob. as } N \rightarrow \infty$$

$$\text{ただし } U_N(\theta_0, \hat{\theta}) = \max_{\theta} U_N(\theta_0, \theta)$$

一致性の証明

- Ferguson(1996) (i.i.d. ver.) を異分布ver.に書き換え

$$U_N(\theta_0, \theta) = \frac{1}{N} \log \frac{f(y, \theta)}{f(y, \theta_0)}, \quad \mu_N(\theta_0, \theta) = \frac{1}{N} E_{\theta_0} \frac{f(y, \theta)}{f(y, \theta_0)}$$

$$U_N(\theta_0, \theta) - \mu_N(\theta_0, \theta) \rightarrow 0 \quad \text{Uniformly in prob. as } N \rightarrow \infty$$

$$\forall \varepsilon > 0, \exists C < 0, \underline{\sup_{\theta \in S} \mu_N(\theta_0, \theta) < C} \text{ for } S = \{\theta \in \Theta, \|\theta - \theta_0\| > \varepsilon\}$$

ならば

$$\hat{\theta} \rightarrow \theta_0 \quad \text{in prob. as } N \rightarrow \infty$$

$$\text{ただし, } U_N(\theta_0, \hat{\theta}) = \max_{\theta} U_N(\theta_0, \theta)$$

回帰 $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ の場合

$$\mu_N(\theta_0, \theta) = \frac{1}{2} \left(\log \frac{\sigma_0^2}{\sigma^2} + \left(1 - \frac{\sigma_0^2}{\sigma^2} \right) \right) - \frac{1}{2\sigma^2 N} \|\mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)\|^2$$

$\sup_{\theta \in \mathcal{S}} \mu_N(\theta_0, \theta) < C < 0$ を満たすには

$$\frac{1}{2\sigma^2 N} \|\mathbf{X}(\boldsymbol{\beta} - \boldsymbol{\beta}_0)\|^2 > c_1 > 0.$$

つまり, $\hat{\boldsymbol{\beta}}$ の一貫性には以下の条件が必要.

$$\lambda_{\min} \left(\frac{1}{N} \mathbf{X}^T \mathbf{X} \right) > c_2 > 0$$

- $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$ の形状のみに注目した一貫性なら $\lambda_{\max}((\mathbf{X}^T \mathbf{X})^{-1}) = o(1)$ でよいが, $\lambda_{\max}((\mathbf{X}^T \mathbf{X})^{-1}) = O(1/N)$ でないと, 尤度の安定性は保証できない.

SLM の場合 (独立でも同分布でもないver.)

$$S = \{(\boldsymbol{\beta}, \lambda, \sigma^2): |\boldsymbol{\beta} - \boldsymbol{\beta}_0| > \varepsilon_1 \text{ or } |\lambda - \lambda_0| > \varepsilon_2 \text{ or } |\sigma^2 - \sigma_0^2| > \varepsilon_3\}$$

$$L_A(\boldsymbol{\beta}, \lambda, \sigma^2) = \frac{1}{2} \log \det(A(\lambda, \sigma^2)) - \frac{N}{2} \log 2\pi - \frac{1}{2} \mathbf{u}(\boldsymbol{\beta}, \lambda)^T \tilde{A}(\lambda, \sigma^2) \mathbf{u}(\boldsymbol{\beta}, \lambda)$$

$\boldsymbol{\beta}, \lambda$ の寄与をきれいに分解できない.

そこで, まずは λ と σ^2 のプロフィール尤度を考える.

$$\begin{aligned} L_A(\hat{\boldsymbol{\beta}}, \lambda, \sigma^2) \\ = \frac{1}{2} \log \det(A(\lambda, \sigma^2)) - \frac{N}{2} \log 2\pi - \frac{1}{2} \mathbf{u}^T (I - P(\lambda)) \tilde{A}(\lambda, \sigma^2) (I - P(\lambda)) \mathbf{u} \end{aligned}$$

$$\text{ただし, } P(\lambda) = X(\lambda)(X(\lambda)^T X(\lambda))^{-1} X(\lambda)^T$$

$$2\mu_N(\lambda_0, \lambda) = \frac{1}{N} \sum_k \left\{ \log \frac{d_k(\lambda, \sigma^2)}{d_k(\lambda_0, \sigma_0^2)} + 1 - \frac{d_k(\lambda, \sigma^2)}{d_k(\lambda_0, \sigma_0^2)} \right\}$$

$$\text{ただし, } d_k(\lambda, \sigma^2) = \frac{1}{\sigma^2} |1 - \lambda \cdot \lambda_k(\tilde{W})|^2$$

W が対称なら, 定常性の十分条件

$$|\lambda| < \frac{1}{\max |\lambda_k(\tilde{W})|}$$

だけで (幸いにも) 上からおさえられて, 無条件で一致性が証明できる.

一方, W が非対称の場合, $\lambda_k(\tilde{W})$ が複素数になるため, もう少し複雑になり, まだ証明できていない.

まとめ

- SLMでは漸近バイアスが最悪発散
 - まずは, u の一貫性を確保すべき.
 - L_A のような尤度を使えば漸近有効推定量を得られる.
- 一貫性は当たり前ではない.
 - W が対称ならば追加条件なしでいえる.
 - 非対称なら...??
 - 最尤解はひとつとは限らない.