

回帰項を含む 同時空間自己回帰モデル における最尤推定

早稲田大学
(株) データサイエンスコンソーシアム

力丸佑紀
柴田里程

空間データ解析

◆ Point pattern data analysis →

□ あるイベントが発生する位置 $\{p_i \in R^d\}$

- 例) ある森林内に生えている樹木の位置

イベントの
発生メカニズム

点過程
(Point process)

◆ values on a space analysis →

□ 固定地点で観測された値 $\{X_v, v \in R^d\}$

- 例) 観測所で観測された降雨量データ
- 例) 地価データ
- 例) 画像データ
- 例) 宇宙の温度データ
 - Lattice data $\{X_v, v \in Z^d\}$

値の分布や相関構造

空間過程
(Random field)

空間データの特徴

- ◆ 繰り返し観測ができない
- ◆ 異なる地点の値が互いに影響し合う
 - 「順序」がない

Lattice dataに対するモデル

- ◆ spatial autoregressive model (Whittle, 1954)
 - のちに, simultaneous autoregressive (SAR) model と呼ばれるように.

弱定常過程 $\{Z_v, v \in Z^2\}$ に対し,

$$\sum_{k \in K} \rho_k Z_{v+k} = \varepsilon_v, \quad \varepsilon_v \sim N(\mathbf{0}, \sigma^2)$$

- ε_v は互いに直交する.
- すべての u, v に対して, ε_u は必ずしも Z_v と直交しない.

Whittle のパラメータ推定法

◆ 最尤法

□ 正規分布の仮定のもとでの対数尤度

$$L = -\frac{1}{2} \log \det(\Sigma) - \frac{N}{2} \log 2\pi - \frac{1}{2} \mathbf{z}^T \Sigma^{-1} \mathbf{z}$$

- Σ はパラメータの複雑な関数
- $\det(\Sigma)$ や Σ^{-1} の計算が困難

□ Whittleの近似尤度

$$\frac{1}{4\pi^2} \iint \left[\log f(\omega_1, \omega_2) + \frac{F(\omega_1, \omega_2)}{f(\omega_1, \omega_2)} \right] d\omega_1 d\omega_2$$

$\sqrt{N}$ 一致性を
持たない

- $f(\omega_1, \omega_2)$: スペクトル密度, $F(\omega_1, \omega_2)$: ペリオドグラム

空間データに対するモデル

空間統計学

- ・ 弱定常性を仮定

CARモデルの提案

Besag(1974)など

Whittle近似の改良

Guyon(1982), Robinson
and Sanz(2006)など

Whittle, 1954

空間計量経済学

- ・ 定常性の仮定なし
- ・ 回帰モデルにSAR
の発想を導入

経済学への応用

Anselin(1988), LeSage
and Pace(2006)など

空間計量経済学でのモデル(1)

◆ 回帰モデル $z = X\beta + \epsilon$

↓ 相互依存性を取り入れる

◆ Manskiモデル

$$z = \rho Wz + X\beta + WX\gamma + \epsilon, \quad \epsilon = \lambda W\epsilon + u$$
$$u \sim N(\mathbf{0}, \sigma^2 I)$$

□ 空間重み行列 W

- 空間内の2地点 i, j に依存関係がある場合, それに対応する要素を $w_{i,j} \neq 0$ で与え, 依存関係がない場合, 要素を 0 とする行列



●近接行列

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

●距離の逆数を要素とする行列

$$\begin{pmatrix} 0 & 1 & 1/2 & 1/3 & 1/4 & 1/5 \\ 1 & 0 & 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1 & 0 & 1 & 1/2 & 1/3 \\ 1/3 & 1/2 & 1 & 0 & 1 & 1/2 \\ 1/4 & 1/3 & 1/2 & 1 & 0 & 1 \\ 1/5 & 1/4 & 1/3 & 1/2 & 1 & 0 \end{pmatrix}$$

空間計量経済学でのモデル(2)

◆ 空間ダービンモデル(SDM)

- $\lambda = 0$ のときのManski モデル

$$z = \rho Wz + X\beta + WX\gamma + u, \quad u \sim N(\mathbf{0}, \sigma^2 I)$$

□ 次のように書き換えられる

$$Pz = \varepsilon, \quad \varepsilon = c + u, \quad u \sim N(\mathbf{0}, \sigma^2 I)$$

$$\text{ただし, } P = I - \rho W, \quad c = X\beta + WX\gamma$$

パラメータ推定法

◆ 最尤法

- 計算が困難
 - 計算アルゴリズムの提案
 - 行列式計算の近似
- 推定精度が高い
 - シミュレーション

?

◆ GS2SLS

- 二段階最小二乗法と一般化モーメント法の混合
 - 計算負荷が小さい
 - 推定精度が最尤法ほど高くない

空間ダービンモデルの対数尤度

◆ 空間ダービンモデル

$$P\mathbf{z} = \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} = \mathbf{c} + \mathbf{u}, \quad \mathbf{u} \sim N(\mathbf{0}, \sigma^2 I)$$

$$\text{ただし, } P = I - \rho W, \quad \mathbf{c} = X\boldsymbol{\beta} + WX\boldsymbol{\gamma}$$

正規分布を仮定すると

$$\mathbf{z} \sim N(P^{-1}\mathbf{c}, \sigma^2 P^{-1}(P^{-1})^T)$$

対数尤度は

$$L = \frac{1}{2} \log \det \left(\frac{P^T P}{\sigma^2} \right) - \frac{N}{2} \log(2\pi) - \frac{1}{2\sigma^2} (\mathbf{z} - P^{-1}\mathbf{c})^T P^T P (\mathbf{z} - P^{-1}\mathbf{c})$$

最尤推定がうまくいかない？

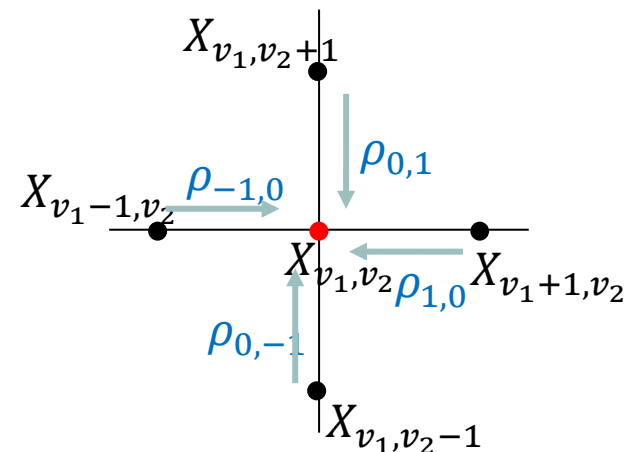
◆ 空間依存性が対称な場合

□ フィッシャー情報量行列が非正則になる (Rikimaru and Shibata, 2017)

● 例) $X_{v_1, v_2} + \rho_{1,0} X_{v_1+1, v_2} + \rho_{-1,0} X_{v_1-1, v_2} + \rho_{0,1} X_{v_1, v_2+1} + \rho_{0,-1} X_{v_1, v_2-1} = \varepsilon_{v_1, v_2}$

真の値が

$\rho_{1,0} = \rho_{-1,0}, \rho_{0,1} = \rho_{0,-1}$
のとき、推定不能になる



◆ 相関がおちない場合

□ ある地点の影響が遠くまで残る

定常過程への近似

- ◆ 空間ダービンモデルは定常性を仮定していない
- ◆ $P = I - \rho W$ がテプリッツ行列（のブロック行列）であれば定常過程に近似できる

弱定常SARモデルの最尤推定

- ◆ 弱定常過程 $\{Z_v, v \in Z^2\}$ に対し,

$$\sum_{k \in K} \rho_k Z_{v+k} = \varepsilon_v, \quad \beta_0 = 1$$

- ◆ 近似対数尤度(Rikimaru and Shibata, 2016)

$$L_A = \frac{1}{2} \log \det(A) - \frac{N}{2} \log 2\pi - \frac{1}{2} z^T \tilde{A} z,$$

$$A = \frac{1}{\sigma^2} \sum_{k, k'} \rho_k \rho_{k'} W_{n_1}^{-k_1+k'_1} \otimes W_{n_2}^{-k_2+k'_2},$$

$$\tilde{A} = \frac{1}{\sigma^2} \sum_{k, k'} \rho_k \rho_{k'} \alpha_{n_1}^{|k_1-k'_1|} \alpha_{n_2}^{|k_2-k'_2|} W_{n_1}^{-k_1+k'_1} \otimes W_{n_2}^{-k_2+k'_2}$$

漸近有効性を持つ
計算も簡単

近似尤度 L_A の特徴

◆ $\{Z_v\}$ から $\{\varepsilon_v\}$ への変換行列を巡回行列で近似

- 変換行列はテプリッツ行列（のブロック行列）
- 変換行列の固有値が求まる
- 変換行列を巡回行列 W_{n_j} を用いて表現できる

$$\sum_{k \in \mathcal{K}} \rho_k W_{n_1}^{k_1} \otimes W_{n_2}^{k_2}$$

◆ 収縮率 $\alpha_{n_j} = 1 + \frac{1}{n_j}$ の導入

- 漸近有効性の成立に必要

巡回行列 W_n の定義と性質

$n \times n$ の巡回行列 $W_n =$

$$\begin{pmatrix} 0 & 1 & & 0 \\ 0 & 0 & 1 & \\ & \ddots & \ddots & \ddots \\ 0 & & 0 & 1 \\ 1 & & & 0 \end{pmatrix}$$

$$W_n^2 = \begin{pmatrix} & & & 0 \\ & & & \\ & & & \\ & & & \\ 0 & & & \\ & & & \\ & & & \\ & & & \end{pmatrix}$$

$$W_n^3 = \begin{pmatrix} & & & 0 \\ & & & \\ & & & \\ & & & \\ & & & \\ 0 & & & \\ & & & \\ & & & \end{pmatrix}$$

特に, $W_n^0 = I_n$, $W_n^{-1} = W_n^T$ である.

$\otimes$ はクロネッカー積

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & & & \\ \vdots & \ddots & \ddots & \vdots \\ a_{n1}B & \cdots & \cdots & a_{nn}B \end{pmatrix}$$

空間ダービンモデルの近似尤度

◆ d次元空間ダービンモデル

$$P\mathbf{z} = \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} = \mathbf{c} + \mathbf{u}, \quad \mathbf{u} \sim N(\mathbf{0}, \sigma^2 I)$$

□ 対数尤度は

$$L = \frac{1}{2} \log \det \left(\frac{P^T P}{\sigma^2} \right) - \frac{N}{2} \log(2\pi) - \frac{1}{2\sigma^2} (\mathbf{z} - P^{-1}\mathbf{c})^T P^T P (\mathbf{z} - P^{-1}\mathbf{c})$$



P を巡回行列B に置き換え
二次形式の項には収縮率を導入

$$L_B = -\frac{1}{2} \log \det \left(\frac{B^T B}{\sigma^2} \right) - \frac{N}{2} \log 2\pi - \frac{1}{2\sigma^2} (\mathbf{z} - B^{-1}\mathbf{c})^T \tilde{B}^T \tilde{B} (\mathbf{z} - B^{-1}\mathbf{c})$$

空間ダービンモデルの近似尤度

- ◆ テプリッツ行列 P を巡回行列 B で書き換え

$$B = \sum_{\mathbf{k} \in \mathcal{K}} \rho_{\mathbf{k}} W_{n_1}^{k_1} \otimes \cdots \otimes W_{n_d}^{k_d}$$

- $-\rho W$ の要素を $\rho_{\mathbf{k}}$ として定義

- ◆ 収縮率を導入した行列 $\tilde{B}$

$$\tilde{B} = \sum_{\mathbf{k} \in \mathcal{K}} \rho_{\mathbf{k}} \prod_{j=1}^d \alpha_j W_{n_1}^{k_1} \otimes \cdots \otimes W_{n_d}^{k_d}$$

$$\text{ただし, } \alpha_j = 1 + \frac{k_j}{n_j} + \frac{c}{n_j^{\frac{d+1}{2}}}$$

- L_A での収縮率 $\left(1 + \frac{1}{n_j}\right)^{|k_j|}$ をそのまま使うと $d \geq 3$ のとき, 漸近有効性が成立しない

空間データに対するモデル

空間統計学

- ・ 弱定常性を仮定

CARモデルの提案

Besag(1974)など

Whittle近似の改良

Guyon(1982), Robinson
and Sanz(2006)など

Whittle, 1954

空間計量経済学

- ・ 定常性の仮定なし
- ・ 回帰モデルにSARの発想を導入

経済学への応用

Anselin(1988), LeSage
and Pace(2006)など