空間自己回帰モデルの特性について

早稲田大学　力丸佑紀（発表者）
（株）データサイエンスコンソーシアム　柴田里程

空間自己回帰(SAR)モデル

空間計量経済学で用いられるモデル

- ２次元空間上の n 個の地域 $\boldsymbol{s}_i (i = 1,2,\cdots,n)$
- 観測値 $Y(\boldsymbol{s}_i)$ / 観測値ベクトル $\boldsymbol{Y}_n$

$$\boldsymbol{Y}_n = \lambda W_n \boldsymbol{Y}_n + X_n \boldsymbol{\beta} + \boldsymbol{\varepsilon}_n$$

  - $X_n \boldsymbol{\beta}$ は回帰項
  - $\boldsymbol{\varepsilon}_n$ は誤差ベクトルで $\boldsymbol{\varepsilon}_n \sim N(\mathbf{0}, \sigma^2 I)$

Lee(2004)

- 最尤法によるパラメータ推定
- 8つの条件が必要

✓ $\boldsymbol{\varepsilon}_n$ は $N(\mathbf{0}, \sigma^2 I)$ にしたがう

✓ W_n の (i,j) 成分 $w_{n,ij}$ は一様に $O(1/h_n)$

✓ $\liminf_{n\to\infty} h_n > 0$ かつ $\lim_{n\to\infty} h_n/n = 0$

✓ $S_n = I - \lambda W_n$ は正則行列

✓ W_n と S_n^{-1} は行和と列和に関して一様に有界

✓ X_n の任意の成分は n に関して一様に有界であり，$\lim_{n\to\infty} X_n^T X_n/n$ が存在して正則行列

✓ Λ を R のコンパクト集合，λ_0 を Λ の内点とする． Λ 上で $S_n^{-1}(\lambda)$ は λ および n に関して一様に行和および列和が有界

✓ $\lim_{n\to\infty} (X_n, G_n X_n \beta_0)^T (X_n, G_n X_n \beta_0)/n$ が存在して，正則行列

どんな条件か？

- 以下の2つの場合について確かめる

1. W_n が隣接行列
   - $\boldsymbol{s}_i$ と $\boldsymbol{s}_j$ が隣り合っていれば $w_{n,ij}=1$，それ以外は 0

2. $Y(\boldsymbol{s}_i)$ が空間オルンシュタイン-ウーレンベック過程の場合
   - $$\mathrm{d}Y_{\boldsymbol{t}} = -\theta_1 X_{\boldsymbol{t}} \mathrm{d}t_1 - \theta_2 X_{\boldsymbol{t}} \mathrm{d}t_2 + \sigma \mathrm{d}B_{\boldsymbol{t}}, \quad t_1, t_2 > 0,\ \ \theta_1 + \theta_2 > 0$$
   - $B_{\boldsymbol{t}}$ は Brownian sheet

1. W_n が隣接行列

- 地域 $\boldsymbol{s}_i$ は格子点上に並んでいるとする
  - $n_1 \times n_2 = n$ の観測データ
  - $\boldsymbol{Y}_n$ は座標が辞書式になるように並べる

- $X_n\boldsymbol{\beta} = 0$ の場合を考える

$$\boldsymbol{Y}_n = \lambda W_n \boldsymbol{Y}_n + \boldsymbol{\varepsilon}_n$$

- $W_n = \begin{pmatrix} A & I & & \\ I & \ddots & \ddots & \\ & \ddots & \ddots & I \\ & & I & A \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 0 \end{pmatrix}$

ブロック行列

確認すべき条件

1. W_n の (i,j) 成分 $w_{n,ij}$ は一様に $O(1/h_n)$
2. $\liminf_{n\to\infty} h_n > 0$ かつ $\lim_{n\to\infty} h_n/n = 0$
3. $S_n = I - \lambda W_n$ は正則行列
4. W_n と S_n^{-1} は行和と列和に関して一様に有界
5. Λ を R のコンパクト集合，λ_0 を Λ の内点とする． Λ 上で $S_n^{-1}(\lambda)$ は λ および n に関して一様に行和および列和が有界

条件 1, 2

1. W_n の (i,j) 成分 $w_{n,ij}$ は一様に $O(1/h_n)$
2. $\liminf_{n\to\infty} h_n > 0$ かつ $\lim_{n\to\infty} h_n/n = 0$

- W_n の (i,j) 成分 $w_{n,ij}$ は 0 または 1
- h_n を定数にとれば条件 1, 2 は成立

条件 3

3.　$S_n = I - \lambda W_n$ は正則行列

- $|S_n| = |I - \lambda W_n| \neq 0$

$\Leftrightarrow$ $1/\lambda$ は W_n の固有値ではない

- $\xi_{j_1,j_2}(W_n) = 2\left\{\cos\left(\frac{j_1\pi}{n_1+1}\right) + \cos\left(\frac{j_2\pi}{n_2+1}\right)\right\},$

$$j_1 = 1,2,\dots,n_1,\ \ j_2 = 1,2,\dots,n_2$$

- $n \to \infty$ のとき $\xi_{j_1,j_2}(W_n)$ は $(-4,4)$ を満遍なく埋め尽くす

- よって，S_n の正則条件は $|\lambda| \leq 1/4$

条件 4, 5

4. W_n と S_n^{-1} は行和と列和に関して一様に有界
5. Λ を R のコンパクト集合，λ_0 を Λ の内点とする． Λ 上で $S_n^{-1}(\lambda)$ は λ および n に関して一様に行和および列和が有界

- 任意の n, i, j に対して，W_n の j 列目の行和は

$$\sum_{i=1}^{n} |w_{n,ij}| = 2\ or\ 3\ or\ 4$$

- W_n の列和は一様に有界
- 列和についても同じ

条件 4, 5

4. W_n と S_n^{-1} は行和と列和に関して一様に有界
5. Λ を R のコンパクト集合，λ_0 を Λ の内点とする．Λ 上で $S_n^{-1}(\lambda)$ は λ および n に関して一様に行和および列和が有界

- S_n^{-1} を固有値分解

$$S_n^{-1} = UD^{-1}U^T$$

  - U は S_n の固有ベクトルを並べた行列
  - D は S_n の固有値を並べた対角行列
- j 列目の要素の絶対値の和は

$$\boldsymbol{v}^T S_n^{-1} \boldsymbol{e}_j = \boldsymbol{v}^T UDU^T \boldsymbol{e}_j = \sum_\ell d_\ell^{-1} \boldsymbol{v}^T \boldsymbol{u}_\ell u_{j\ell}$$

$$\leq \max |d_\ell^{-1}| \sum_\ell u_{j\ell}^2 = \max |d_\ell^{-1}|$$

条件 4, 5

- $S_n = I - \lambda W_n$ より

$$d_{j_1,j_2}(S_n) = 1 - 2\lambda\left\{\cos\left(\frac{j_1\pi}{n_1+1}\right) + \cos\left(\frac{j_2\pi}{n_2+1}\right)\right\}$$

- $|\lambda| < 1/4$ のとき，S_n^{-1} の行和が一様に有界
  - 条件 3 を満たす条件は $|\lambda| \leq 1/4$ だった

- 列和に関しても同じ

空間自己回帰(SAR)モデル

- 空間計量経済学で用いられるSARモデル

$$\boldsymbol{Y}_n = \lambda W_n \boldsymbol{Y}_n + \varepsilon_n$$

  - W_n は隣接行列

- Lee(2004)の最尤法によるパラメータ推定に必要な条件は

$$|\lambda| < 1/4$$

- これは弱定常性の条件と同じ条件

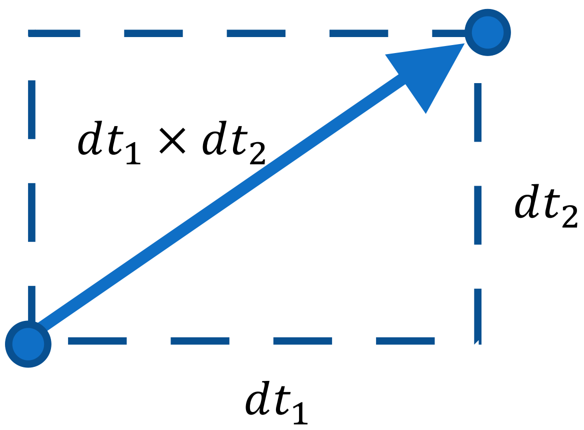
$Y(\boldsymbol{s}_i)$ が空間O-U過程

- 地域 $\boldsymbol{s}_i$ は格子点上に並んでいるとする
  - $n_1 \times n_2 = n$ の観測データ

- $Y(\boldsymbol{s}_i)$ が空間オルンシュタイン-ウーレンベック過程の場合
  - $\mathrm{d}Y_{\boldsymbol{t}} = -\theta_1 Y_{\boldsymbol{t}} \mathrm{d}t_1 - \theta_2 Y_{\boldsymbol{t}} \mathrm{d}t_2 + \sigma \mathrm{d}B_{\boldsymbol{t}},$
    $$t_1, t_2 > 0, \quad \theta_1 + \theta_2 > 0$$
  - $B_{\boldsymbol{t}}$ は Brownian sheet

Brownian sheet

- John Walsh(1975)
- a finitely additive set function W defined on the Borel subsets of $\boldsymbol{R}_{+}^{2}$ such that
  1. $W(A)$ is a $N(0,|A|)$ random variable
  2. if $A \cap B = \emptyset$, then $W(A)$ and $W(B)$ are independent



SARモデルへの書き換え

- $\mathrm{d}Y_{\boldsymbol{t}} = -\theta_1 Y_{\boldsymbol{t}} \mathrm{d}t_1 - \theta_2 Y_{\boldsymbol{t}} \mathrm{d}t_2 + \sigma \mathrm{d}B_{\boldsymbol{t}}$ の解は

$$Y_{\boldsymbol{t}} = \sigma B\left(\frac{e^{2\theta_1 t_1}}{2\theta_1}, \frac{e^{2\theta_2 t_2}}{2\theta_2}\right) e^{-\boldsymbol{\theta}^T \boldsymbol{t}}$$

- これを用いると自己回帰表現が可能

$$Y(\boldsymbol{t}) = e^{-\boldsymbol{\theta}^T \boldsymbol{1}} Y(\boldsymbol{t} + \boldsymbol{1}) + \varepsilon(\boldsymbol{t})$$

微分方程式の解

- $z_{\boldsymbol{t}} = e^{\boldsymbol{\theta}^T \boldsymbol{t}} y_{\boldsymbol{t}}$ とおくと，伊藤の公式より，
$$Y_{\boldsymbol{t}} = e^{-\boldsymbol{\theta}^T \boldsymbol{t}} Y_{\boldsymbol{0}} + \sigma \int_{\boldsymbol{0}}^{\boldsymbol{t}} e^{-\boldsymbol{\theta}^T (\boldsymbol{t}-\boldsymbol{s})} \mathrm{d}B_{\boldsymbol{s}} \quad (1)$$
- また，$e^{\boldsymbol{\theta}^T \boldsymbol{t}} \mathrm{d}B_{\boldsymbol{t}}$ は $\mathrm{d}B\left(\frac{e^{2\theta_1 t_1}}{2\theta_1}, \frac{e^{2\theta_2 t_2}}{2\theta_2}\right)$ と同等 (2)
  - どちらも独立増分であることは確か
  - 分散はどちらも
$$e^{2\boldsymbol{\theta}^T \boldsymbol{t}} dt_1 dt_2$$
- (1)(2) を合わせると前述の解が得られる

SAR表現の導出

- $Y(t) - e^{-\theta^T 1} Y(t-1)$
$$= \sigma e^{-\theta^T t} \left\{ B\left(\frac{e^{2\theta_1 t_1}}{2\theta_1}, \frac{e^{2\theta_2 t_2}}{2\theta_2} \right) - B\left(\frac{e^{2\theta_1 (t_1 - 1)}}{2\theta_1}, \frac{e^{2\theta_2 (t_2 - 1)}}{2\theta_2} \right) \right\}$$
- $Var\left(Y(t) - e^{-\theta^T 1} Y(t-1) \right)$
$$= \sigma^2 e^{-2\theta^T t} \left(\frac{e^{2\theta_1 t_1} (1 - e^{-2\theta_1})}{2\theta_1} \cdot \frac{e^{2\theta_2 t_2} (1 - e^{-2\theta_2})}{2\theta_2} \right)$$
$$= \sigma^2 \cdot \frac{1 - e^{-2\theta_1}}{2\theta_1} \cdot \frac{1 - e^{-2\theta_2}}{2\theta_2}$$
  - t_1, t_2 によらない

定常性

- Brownian sheet について
$$Cov\big(B(\boldsymbol{s}), B(\boldsymbol{t})\big) = \min(s_1, t_1) \cdot \min(s_2, t_2)$$
と定義すると，$Y(\boldsymbol{t})$ の定常性が示せる

W_nのかたち

- 地域 $\boldsymbol{s}_i$ は格子点上に並んでいるとする
  - $n_1 \times n_2 = n$ の観測データ
- $X_n\boldsymbol{\beta} = 0, \lambda = e^{-\boldsymbol{\theta}^T\mathbf{1}}$ の場合のSARモデル

$$\boldsymbol{Y}_n = e^{-\boldsymbol{\theta}^T\mathbf{1}} W_n \boldsymbol{Y}_n + \boldsymbol{\varepsilon}_n$$

  - $W_n = \begin{pmatrix} O & B & & \\ & \ddots & \ddots & \\ & & \ddots & B \\ & & & O \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}$

確認すべき条件

1. W_n の (i,j) 成分 $w_{n,ij}$ は一様に $O(1/h_n)$
2. $\liminf_{n\to\infty} h_n > 0$ かつ $\lim_{n\to\infty} h_n/n = 0$
3. $S_n = I - \lambda W_n$ は正則行列
4. W_n と S_n^{-1} は行和と列和に関して一様に有界
5. Λ を R のコンパクト集合，λ_0 を Λ の内点とする． Λ 上で $S_n^{-1}(\lambda)$ は λ および n に関して一様に行和および列和が有界

条件の確認

- 条件 1, 2 は h_n を定数におけばよい
- S_n は対角成分がすべて 1 の上三角行列なので常に$|S_n| = 1$で正則（条件 3）
- W_n の行和と列和は一様に有界（条件4）
- $S_n^{-1} = I + e^{-\boldsymbol{\theta}^T \mathbf{1}} W_n$ なので，S_n^{-1} の行和または列和は 1 または $1 + e^{-\boldsymbol{\theta}^T \mathbf{1}}$ なので一様に有界（条件 4, 5）

- $Y(\boldsymbol{t})$が空間 O-U 過程の場合，Lee(2004)の条件は常に満たされる

Leeの条件は何を意味するか？

- $X_n \boldsymbol{\beta} = 0$ の場合には，定常性を仮定していることと同じ？？？
  - 特に S_n に関する条件…
  - 今回は W_n に関する条件はきいていない