空間データセットにおける内生的な空間相関構造の同定に関する考察*

Identification of Endogenous Spatial Dependence Structure from a Spatial Data Set *

塚井　誠人**・奥村　誠***

By Makoto TSUKAI **・ Makoto OKUMURA ***

１．はじめに

空間計量経済モデルは，外生変数を含む構造方程式と誤差の空間相関構造を特定化したモデルであり，構造方程式に含まれるパラメータと共に空間相関パラメータを設定する．ただし空間計量経済モデルでは，ARIMAモデルのようにラグごとにパラメータを設定することはできない．通常は，自己回帰型や移動平均型などによって誤差の空間相関構造を同定した上で，各地点の相対的な空間相関構造を空間重み付け行列として設定し，空間相関パラメータがその重みを調整する役割を担う．

換言すると空間相関パラメータは，所与の誤差相関構造と空間重み付け行列の下で，空間相関のスケールを決定する役割を果たす．空間重み付け行列と距離や所要時間等の各近接性指標とは，それぞれが一対一で対応する．つまり空間計量経済モデルを推定する場合，候補となる空間重み付け行列（＝近接性指標）と誤差相関構造の中から，最適なモデルを選択しなくてはならない．

しかし実証分析において，候補となる複数の近接性指標から最も適切な空間重み付け行列を選択することは容易ではない．これは多くの近接性指標が，基本的には地点間距離を反映していて各指標間の実質的な違いが小さく，モデル間に有意差が生じにくいためである．また,複数の近接性指標を任意のパラメータで加重した合成指標を用いて空間重み付け行列を設定すると，推定上は外生変数となる加重パラメータに関する統計量が求められない．そのため，適切な空間重み付け行列の選択はさらに難しくなる．このように空間データが示す空間相関特性を捉える上で，唯一の近接性指標に基づいて設定した空間重み付け行列の下で，ただ１つの空間相関パラメータを推定する空間計量経済モデルは，モデル構造が要求する制約が強く，柔軟性に欠ける点がある．したがって空間データの空間相関特性をより柔軟に表現するには,多属性の近接性指標や地域特性を反映した空間相関構造の設定ができる方が望ましい．

本研究では，空間相関パラメータを地域間特性を表わす多属性ベクトルの関数（以下，空間相関スケール関数）として定式化した空間計量経済モデルを提案する．提案したモデルでは，空間相関スケール関数に含まれるハイパーパラメータを推計することによって，各ハイパーパラメータの影響を統計的に検定することができる．また推定されたハイパーパラメータと空間重み付け行列から得られる空間相関構造を観察することによって，その特性を理解することができる．

*キーワーズ：観光交通，宿泊施設立地，日帰り交通圏

**正員，博（工），広島大学大学院工学研究科
（広島県東広島市鏡山１－４－１
TEL082-424-7827，FAX082-424-7827）
***正員，博（工），東北大学東北アジア研究センター
（宮城県仙台市青葉区川内４１
TEL022-795-7571，FAX022-795-7477）

２．空間相関構造と空間重み付け行列の設定

空間計量経済モデルには，誤差相関関数の設定に関して，いくつかのバリエーションがある[1)]．以下，式（1）に示すSARモデルについて考えよう．

$$\mathbf{Y} = \rho \mathbf{W}\mathbf{Y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

ここで地点数をN，外生変数の数をKとすれば，$\mathbf{Y}$は内生変数（$N \times 1$），$\mathbf{X}$は外生変数（$N \times K$），$\mathbf{W}$は空間重み付け行列（$N \times N$），$\boldsymbol{\varepsilon}$は各地点間で独立かつ同一の分布にしたがう誤差項である．$\boldsymbol{\beta}$は構造方程式のパラメータ，ρはスカラーの空間相関パラメータである．ただしカッコ内はベクトルのサイズを表わす．SARモデルの空間相関構造は,

$$\mathbf{Y} = \mathbf{A}^{-1}\mathbf{X}\boldsymbol{\beta} + \mathbf{A}^{-1}\boldsymbol{\varepsilon} \tag{2}$$

となる（ただし，$\mathbf{A} = \mathbf{I} - \rho \mathbf{W}$）．誤差項は$\mathbf{A}^{-1}\boldsymbol{\varepsilon}$であるので，地点間で独立ではない．このようにSARモデルでは地点iの誤差項は他地点jと相関しており，かつ空間重み付け行列$\mathbf{W}$の設定が空間相関構造を左右する．

空間重み付け行列$\mathbf{W}$として多く用いられるのは，地域ij間の隣接の有無をダミー変数で表現した隣接行列,地域間の距離d_{ij}の逆数等に基づいて隣接の程度を定義した距離行列などである．式(1)から明らかなように,空間相関パラメータρは$\mathbf{W}$の観測スケールの影響を受ける．したがって，隣接行列や距離行列によって$\mathbf{W}$を

定義すると，モデル間でρの推定値が比較できない．後述するように，空間確率過程の定常性条件としてρに課せられる制約条件も，$\mathbf{W}$のスケールによって変化する．そこで実証分析では，スケールを基準化した空間重み付け行列が用いられる．代表的な行基準化（row-standardization）を式（3）に示す．

$$w_{ij} = \frac{d_{ij}^{-\alpha}}{\sum_j d_{ij}^{-\alpha}} \quad (3)$$

ここでw_{ij}は$\mathbf{W}$の要素，d_{ij}は地域ij間の近接性指標，および$\alpha > 0$は空間減衰パラメータである．式（3）より行基準化を行うと，

$$\sum_j w_{ij} = 1, \forall i \quad (4)$$

が成立する．なお$\mathbf{W}$の基準化方法も空間相関構造を左右する．この点に関する議論はTiefelsdorf[2]を参照して欲しい．

式（3）に現れる距離減衰パラメータαは，1また2などの値が外生的に設定される[3]．しかし式（3）が示すようにαも空間相関構造を左右する．αについては，地球統計学の分野で開発されたKriging法の観点から，批判的な検討が進められている．Kriging法は，対象空間内の複数の既知観測点の空間多変量分布から，対象空間の内挿点の未知変量予測を，精度良く行うこと目的として開発された手法である[4]．Kriging法では，空間重み付け行列$\mathbf{W}$の代わりに，地点間距離d_{ij}の関数として，共分散構造$K\left(d_{ij}\right)$を定式化する．SARモデルの共分散関数Kは，

$$\mathbf{K} = \left[\left(\mathbf{I} - \rho\mathbf{W}\right)'\left(\mathbf{I} - \rho\mathbf{W}\right)\right] \quad (5)$$

である．代表的な共分散関数を以下に示す．

$$K\left(d_{ij}\right) = b_1 \exp\left(-b_2 d_{ij}\right) \ \left(d_{ij} > 0\right) \quad (6)$$

$$K\left(d_{ij}\right) = b_1 \exp\left(-b_2 d_{ij}\right) \ \left(d_{ij} > 0\right) \quad (7)$$

$$K\left(d_{ij}\right) = b_1 \left(1 - \frac{3d_{ij}}{2b_2} + \frac{3d_{ij}^3}{2b_2^3}\right) \ \left(0 < d_{ij} < b_2\right) \quad (8)$$

ただし$b_1, b_2 > 0$，$K(0) = 1$である．なお空間内挿を行うKriging法では，行基準化を行うと地点毎に距離の計量が変化するため，式（3）のような基準化は行わない．

堤らは，Kriging法の観点からSARモデルやSMAモデルを検討し，これらのモデルが分散均一性や等方性（同一距離地点間の共分散が等しい性質）を満たさないことを指摘した[5) 6)]．一連の研究では，不規則に観測点が分布する対象空間について，観測点と非観測点を含む母空間を十分に多数の点を含む空間で近似した上で，非観測点から観測点への空間相関の影響を考慮したモデルを定式化して，Kriging法と整合的な空間計量経済モデルを定式化している．Dubin[7]は，SARモデルに移動平均誤差を付加したSARMAモデルにおいて，式（6）～（8）のKriging型重み付け行列を設定し，モンテカルロシミュレーションによってモデルの特定化の誤りが構造パラメータの推定に及ぼす影響を検討した．さらに瀬谷ら[8]は，同様のアプローチを時空間データに拡張したモデルを提案し，実証研究を試みている．このようにKriging型重み付け行列を用いれば，αの設定に関する恣意性を回避することができる上，通常の空間計量経済モデルの設定よりもデータが示す空間相関の減衰特性を柔軟に表現できる．

社会経済的な変数を扱う空間計量経済モデルでは，内生変数や誤差の空間相関構造を柔軟に表現するため，距離以外の所要時間や一般化費用などの近接指標に基づく空間重み付け行列が用いられる[9]．実証分析では，空間重み付け行列と空間相関パラメータを複数組設定した，Multi parametric modelが提案されている．式（1）において，空間重み付け行列$\mathbf{W}_1$，$\mathbf{W}_2$を設定した，SAR型のMulti parametric modelを示す．

$$\mathbf{Y} = \rho_1 \mathbf{W}_1 \mathbf{Y} + \rho_2 \mathbf{W}_2 \mathbf{Y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (9)$$

ここでρ_1, ρ_2は，それぞれ$\mathbf{W}_1$，$\mathbf{W}_2$に対応した空間相関パラメータである．Multi parametric modelを用いた実証分析の例は，Rietveld, and Wintershoven[10]，および塚井ら[11]に見られる．

３．Hyper Parametric Modelの提案

(1)最尤推定量とρへの制約条件

式（1）において，誤差項$\boldsymbol{\varepsilon}$に正規分布$N(0, \sigma^2\mathbf{I})$を仮定し，パラメータベクトルを$\xi = (\rho, \beta, \sigma)$とすれば，SARモデルの対数尤度関数（10）が得られる．

$$\begin{aligned} L\left(\xi \mid \mathbf{X}, \mathbf{Y}\right) &= -\frac{n}{2}\log\left(2\pi\right) + \frac{n}{2}\log\left(\sigma^2\right) + \log\left|\mathbf{A}\right| \\ &= \sigma^{-2}\left(\mathbf{AY} - \mathbf{X}\boldsymbol{\beta}\right)'\left(\mathbf{AY} - \mathbf{X}\boldsymbol{\beta}\right) \end{aligned} \quad (10)$$

対数尤度関数（10）には，誤差項$\boldsymbol{\varepsilon}$の空間相関を表わす$\mathbf{A} = \mathbf{I} - \rho\mathbf{W}$のdeterminantの対数項$\log\left|\mathbf{A}\right|$が含まれている．よって最尤推定量$\hat{\xi}_{ML} = \arg\max L\left(\xi \mid \mathbf{X}, \mathbf{Y}\right)$が実数であるには，$\left|\mathbf{A}\right| > 0$が成立しなければならない．

ここで$\mathbf{A}$を，式(3)によって行方向に基準化された空間重み行列$\mathbf{W}$としよう．$\mathbf{A}$は非対称なN次正方行列である．$\mathbf{A}$の2次形式$\mathbf{x}'\mathbf{A}\mathbf{x}$について，以下の恒等式が成立する[12]．

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{A}'\mathbf{x} = \mathbf{x}'\left(\frac{\mathbf{A} + \mathbf{A}'}{2}\right)\mathbf{x} = \mathbf{x}'\mathbf{B}\mathbf{x} \quad (11)$$

式（11）より，$\mathbf{B}$の正値定符号条件を求めれば，$\mathbf{A}$の

正値定符号条件が得られることがわかる．$\mathbf{A}$ の定義に基づいて $\mathbf{B}$ の要素を検討すると，$\mathbf{B}$ の対角要素 b_{ii}，および非対角要素 $b_{ij}, j \neq i$ は，それぞれ，$b_{ii} = 1$，$b_{ij} = b_{ji} = -(w_{ij} + w_{ji})/2$ である．w_{ij}，w_{ji} はそれぞれ行基準化されているので，$(w_{ij} + w_{ji})/2 < 1$ である．

以上より $\mathbf{B}$ は対角項が1で，ρ の係数の絶対値は1より小さい．これは相関行列と同一の構造である．相関行列の正値定符号条件は，$\mathrm{rank}\,\mathbf{B} = N$，および $|b_{ij}| < 1$ である．このうち前者は，データセットに地域特性が同一のサンプルが複数含まれていない限り，通常は満たされる．行基準化された空間重み行列 $\mathbf{W}$ の要素 w_{ij} の絶対値は，たとえば地域 i の隣接地域が1つの場合の接続行列などでは1の可能性があることに注意すると，最大で1をとる．したがってパラメータ ρ に関する実質的な制約条件は，$|\rho| < 1$ である．この条件は，空間確率過程の定常性条件と一致する．

ここでKriging法の共分散関数とMulti parametricモデルの共分散関数についても，正値定符号条件を確認しておこう．式（6）～（8）の b_1, b_2 の最尤推定量の実数条件は，空間確率過程の定常性条件と一致し，共分散関数 $K(d)$ に正値定符号条件（$b_1, b_2 > 0$）が課せられる．式（9）の ρ_1, ρ_2 については，$\mathbf{A} = |\mathbf{I} - \rho_1\mathbf{W_1} - \rho_2\mathbf{W_2}|$ である．尤度関数（10）について検討すると上述の議論と同様に，$\mathbf{W} = \mathbf{W}_1 + \mathbf{W}_2$ かつ $\mathbf{W}$ が行基準化されている場合の最尤推定量の実数条件は，ρ_1, ρ_2 がそれぞれ独立に $|\rho_1| < 1$，および $|\rho_2| < 1$ を満たすことである．これは $\rho_1 - \rho_2$ 平面では ρ_1, ρ_2 が原点対称な矩形領域の内部になくてはならないことを意味する．一方 $\mathbf{W}_1$，$\mathbf{W}_2$ がそれぞれ基準化されている場合は，２つの空間相関パラメータの最尤推定量が実数条件を満たす範囲は独立ではなく，$\rho_1 - \rho_2$ 平面で ρ_1, ρ_2 が条件を満たす範囲は矩形領域とはならない．この領域の形状は空間重み付け行列に左右されるため，解析的に求めることはできない[13]．

前者のMulti parametricモデルでは $\mathbf{W} = \mathbf{W}_1 + \mathbf{W}_2$，であるので，$\mathbf{W}$ を分割して $\mathbf{W}_1, \mathbf{W}_2$ とする基準，たとえば地域 ij を相互に排反な集合に分割するなどの設定の妥当性が問題となる．一方後者のMulti parametricモデルではそのような問題は生じないが，最尤推定量の実数条件を満たす空間相関パラメータの範囲が複雑になってしまう．この場合，類似した近接性指標から設定した空間重み付け行列を複数用いると多重共線性が生じるなど，空間相関パラメータを安定的に推計しにくくなる．

(2)Hyper Parametric SARモデル

3(1)において，空間重み行列のパラメータ ρ には条件 $|\rho| < 1$ が課せられることを確認した．導出の過程によれば，条件 $|\rho_{ij}| < 1$ を満たしていれば，地域 ij ごとに異なる ρ_{ij} を設定しても良い．しかし空間重み付け行列の非対角項全てに，自由な $N(N-1)$ 個の空間相関パラメータ ρ_{ij} を設定すると，サンプル数 N に対して，自由度が不足する．

そこで ρ_{ij} を特性ベクトル

$$z_{ij} = \sum_k \gamma^k u_{ij}^k \tag{12}$$

の関数として設定し，$\mathbf{z}$ に含まれるハイパーパラメータ γ^k を構造方程式のパラメータと共に推定することによって自由度を確保し，かつ地域間の特性を反映した空間相関関係を表現しよう．条件 $|\rho_{ij}| < 1$ を満たす連続かつ２回微分可能な関数として，双曲線正接関数（13）を仮定する．

$$\rho(z_{ij}) = \frac{2}{1+\exp(-z_{ij})} - 1 = \frac{1-\exp(-z_{ij})}{1+\exp(-z_{ij})} = \tanh(-z_{ij}) \tag{13}$$

式(13)から明らかなように，双曲線正接関数はロジスティック曲線のアフィン変換によって得られる．したがって，これらの関数は幾何的に相似である．双曲線正接関数の定義域は $-\infty < z_{ij} < \infty$，値域は $-1 < \rho(z_{ij}) < 1$ である．なお，特性ベクトル $\mathbf{z}$ は線形関数であるので，$\gamma^k = 0$ のとき $\rho(z_{ij}) = 0, \forall i, j$ である．すなわち帰無仮説 $H_0 : \gamma^k = 0$ の仮説検定を行えば，説明変数 $\mathbf{u^k}$ が空間相関に及ぼす影響が検定できる．

以上の考察より，双曲線正接関数（13）を用いれば，制約条件 $|\rho_{ij}| < 1$ を満たしつつ，特性ベクトル $\mathbf{z}$ とそのハイパーパラメータベクトル $\boldsymbol{\gamma}$ によって，より柔軟に空間相関パラメータ $\rho(z_{ij})$ を定式化できる．地域間 ij の特性ベクトル $\mathbf{z}$ に応じて空間的な依存関係が異なる空間自己回帰モデルを，以下H-SAR:Hyper parametric SARモデルと呼ぶ．

$$\mathbf{Y} = \rho(\mathbf{z}) \otimes \mathbf{WY} + \mathbf{X\beta} + \boldsymbol{\varepsilon} \tag{14}$$

ここで $\mathbf{z}$ は式（12）に示す要素 z_{ij} を持ち，対角要素について $z_{ii} = 0$ を満たす N 次正方行列である．また記号 $\otimes$ は要素積（クロネッカー積）を表わす．

４．Hypar Parametric SARモデルの適用

(1)業務立地量モデルの定式化

時点 t，地域 i の業務立地量を E_i^t，時点 t，地域 $j \neq i$ の業務立地量を E_i^t，および時点 $t-1$，地域 i の業務立地量を E_i^{t-1} とする．ここで，業務立地量モデルとして，式（15）を仮定する．

$$\Delta E_i^t = \sum_j^{j \neq i} \rho(z_{ij}^t) w_{ij} \Delta E_j^t + \beta_1 + \beta_2 E_i^{t-1} + \varepsilon_i \tag{15}$$

左辺の内生変数はt時点の業務立地量E_i^tと，$t-1$時点の業務立地量E_i^{t-1}の差$\Delta E_i^t = E_i^t - E_i^{t-1}$である．一方右辺は，空間重み係数$\rho\left(z_{ij}^t\right)w_{ij}$によって加重された$t$時点の他地域$j \neq i$の業務立地変化量$\Delta E_i^t$と前時点業務立地量$E_i^{t-1}$，および誤差項$\varepsilon_i$である．推定された空間依存関係$\rho\left(z_{ij}^t\right)w_{ij}$を観察することにより，地域間の補完・競合関係を内生的に求めることができる．式（15）は，ε_iに正規分布を仮定するとHyper parametric型のSARモデルとなる．パラメータβ, σ，およびγ^kは最尤法によって同時推定する．

(2)データおよび推定結果

分析対象地域は，幹線旅客純流動調査で用いられた全国207生活圏のうち，沖縄と離島を除く194地域とする．幹線交通ネットワークは，194地域に対応させて作成したネットワーク（240ノード，501リンク）上で，2000年の時刻表データに基づいて作成したデータベースを用いて，最短経路探索を行った．なお最低運行頻度は，最短経路上で最も運行頻度の低いリンクの1日あたりの運行頻度である．空間重み行列$\mathbf{W}$の各要素w_{ij}は，各生活圏間の代表点間の距離に基づいて算出し，行基準化を行った．各地域の代表地点を地域内で最も人口の多い各市町村役場の所在地として，国土数値情報（公共施設（点））[14]に基づく緯度経度から測地線距離を算出した．また業務立地量データとその他の地域特性指標は，民力CD-ROM2004（朝日新聞社）より，1990年と2000年の市町村別の事業所立地数を194地域について集計した．

表-1にモデルの推定結果を示す（カッコ内はｔ値）．なお推定結果の詳細は発表時に述べる．

表-1　H-SAR型事業所立地モデルの推定結果

説明変数	H-SAR	SAR	OLS
定数項	371.3	-285.2	655.9
	(0.59)	-(0.48)	(3.17)
90年立地量	-0.086	-0.085	-0.084
	-(26.75)	-(26.75)	-(26.63)
所要時間逆数	0.104	-	-
	(2.04)	-	-
最低運行頻度	-0.340	-	-
	(2.21)	-	-
二次産業比	-20.00	-	-
	(2.97)	-	-
ｚｉｊ定数項	5.510	-	-
	(2.82)	-	-
ρ	-	-0.205	-
	-	-(1.70)	-
分散	2267.4	2448.5	2480.1
	(19.67)	(19.70)	-
最終尤度	-1775.3	-1789.1	-1790.6
AIC	3538.6	3570.3	3577.2
Adjusted R2	0.757	0.783	0.787
Moran's I	0.556	1.783	6.079
サンプル数		194	

４．おわりに

本論では，空間データセットに基づいて，内生的に空間相関構造を推定することができるモデルとして，Hyper Parametric型の空間相関構造を持つ空間計量経済モデルを提案し，その適用例を示した．

参考文献

１）Anselin, L.： Spatial econometrics: methods and models, Kluwer academic publishers, 1988.
２）Tiefelsdorf, M.：Modelling spatial process, Springer, 2000.
３）古谷知之，原田昇，太田勝敏：空間統計モデルを用いた都心内居住特性と就業人口予測に関する研究, 都市計画論文集, No.35, pp.247-252, 1999.
４）Wackernagel, H.（地球統計学研究委員会訳）：Geostatistics（邦題：地球統計学）,森北出版,2003.
５）堤盛人，清水英範，井出裕史：空間的自己相関を記述するための重み行列の構造が分析結果に及ぼす影響，土木計画学研究・論文集，No.17，pp.321-325, 2000.
６）堤盛人，清水英範，井出祐史：移動平均モデルに基づく Kriging を用いた空間内挿，土木計画学研究・講演集，No.23(2)，pp.483-486，2000.
７）Dubin, R.: Spatial lags and spatial errors revisited: Some Monte Carlo evidence, Spatial and spatiotemporal econometrics (Advances in econometrics vol.18), pp.75-98, 2004.
８）瀬谷創，堤盛人，井上亮，石田東生，岡本直久：Covariogram に移動平均モデルを用いた時空間Kriging，第２０回応用地域学会研究発表会発表論文，2006.
９）Bao, S., Henry, M. and Barkley, D.：Identifying urban-rural linkages: Tests for spatial effects in the Carlino-Mills model, Advances in spatial econometrics, Springer, pp.321-334, 2004.
10）Rietveld, P. and Wintershoven, P.: Border effects and spatial autocorrelation in the supply of network infrastructure, Papers in regional science, vol.3, pp.265-276, 1998.
11）塚井誠人，江尻良，奥村誠，小林潔司：社会資本の生産性とスピルオーバー効果，土木学会論文集, No.716/IV-57,pp.53-67, 2002.
12）Schott, J.：Matrix analysis for statistics, second edition, Wiley & Sons, 2004.
13）Hepple, L.: Bayesian model choice in spatial econometrics, Spatial and spatiotemporal econometrics (Advances in econometrics vol.18), pp.101-126, 2004.
14）http://nlftp.mlit.go.jp/ksj/(2006.11)