空間的自己相関論(II)：ノイズ効果とその除去

著者	奥野 隆史
雑誌名	筑波大学人文地理学研究
巻	12
ページ	1-24
発行年	1988-03-25
URL	http://hdl.handle.net/2241/00151084

人文地理学究研　XII　1〜24　1988

空間的自己相関論（II）
―ノイズ効果とその除去―

奥野隆史

I　序　　論
II　空間的自己相関の既存統計量に対する影響
　1.　t 統計量について
　2.　回帰パラメータについて
III　空間的自己相関の除去
　1.　減算法
　2.　修正 t 統計量
　3.　一般化最小2乗法
　4.　最尤法

I　序　　論

前報告において，筆者は，空間的自己相関の問題が計量地理学における基礎的かつ最も重要な問題の1つであり，さらにそれの研究が地理学固有の課題である地域性の解明に係わっていることを指摘した（奥野，1981）．そして，各種の空間データにおける空間的自己相関の有無と強度を検出するための諸測度について，従来の多数の研究に基づいて説述した．

この空間的自己相関問題に関する研究の系譜は，田中（1982）によって指摘されたように，概略的には，計量経済学における時系列自己相関研究および計量植物生態学における分布生成過程研究と深く関係しながら，初めに，空間的自己相関の存在に対する認識とその存在の検出のための測度の定式化，次いで，各種の既存統計量に対する自己相関の影響の確認とそれの除去，最近においては自己相関自体のモデル化という潮流を示している．この研究系譜に従うならば，前報告は初期段階の成果をまとめたものといえる．本報告は，それに継続する，空間的自己相関の既存統計量に対する影響――後述のようにこれは明らかにノイズ効果である――とそれの除去について説述する．

II　空間的自己相関の既存統計量に対する影響

われわれは，空間過程を計量的に解明するに際しては，野外観察や聞取りなどをとおして数値データを収集し，それを各種の手法によって分析し，その成果に関する仮説をたて，それを検定するという接近法を取る場合が多い．この仮説の検定は，一般的には統計的推定法の利用によって行なわれる．衆知のように，この推定法は，n 個の互いに独立的に分布するサンプル観測体に関する変数値（属性値）X_i $(i=1, 2, \cdots, n)$ を前提として構築されている．すなわち，観測体を地点とすれば，n 個の地点に関して入手された変数値は互いに独立しているという前提のもとで，統計的推測が行なわれるということである．前報告で述べたように，空間的自己相関は，空間的に隣接した複数の地点の変数値間にみられる相関のことであり，そのような相関を示す変数値は，上述の前提を満たすものではないことは明白である．したがって，空間的自己相関を呈する変数値に対する統計的推測は，その相関の存

在ゆえに何らかの影響を受けることは容易に理解できるであろう[1]．変数値にプラスの空間的自己相関がみられる場合（互いにほぼ同等の変数値が特定の場所に叢集する場合）は，ある1地点の変数値が，その隣接地点のそれによって部分的に説明されるので，その変数値が有する情報は，それが空間的に独立している場合に比べて減少する．したがって，地点が増加すればするほど，情報は全体として減少するであろう．このような空間的自己相関によってもたらされる統計的推測の変動を，従来の諸研究を参照しながら，この推測において頻用される t 統計量および回帰パラメータについて述べる．

1. t 統計量について

Cliff & Ord (1975a) は，空間的自己相関を有する変数値が，互いに独立した変数値を前提として構築されている t 統計量に如何なる偏性をもたらすかについて考察している．

彼らの扱った t 統計量は，2変数それぞれの平均の間の差の有意性検定に利用されるものであり，これは次のように定義されている．$\mathrm{N}(\mu_1, \sigma_1^2\,\mathbf{I})$ なる分布をもつ母集団 X_1 から n_1 個のサンプル値を，そして，$\mathrm{N}(\mu_2, \sigma_2^2\mathbf{I})$ なる分布をもつ母集団 X_2 から n_2 個のサンプル値をそれぞれ抽出し，$\sigma_1^2=\sigma_2^2=\sigma^2$ と仮定すれば，母平均 μ_1 と μ_2 の間には差がないという帰無仮説のもとでは，次の統計量は自由度 (n_1+n_2-2) をもつ t 分布に従う，というのがそれである．

$$t=\frac{\bar{x}_1-\bar{x}_2}{\hat{\sigma}_{\bar{x}_1-\bar{x}_2}} \tag{1}$$

ここでの $\bar{x}_1$ と $\bar{x}_2$ は母集団 X_1 と X_2 から抽出された2サンプル値群それぞれの平均，$\hat{\sigma}_{\bar{x}_1-\bar{x}_2}$ は2サンプル平均間の差の標準誤差の推定値であり，次式によって与えられる．

$$\hat{\sigma}_{\bar{x}_1-\bar{x}_2}=\left(\frac{n_1 S_1^2+n_2 S_2^2}{n_1+n_2-2}\right)^{1/2}\left(\frac{n_1+n_2}{n_1 n_2}\right)^{1/2} \tag{2}$$

ここでの S_1^2 と S_2^2 は2サンプル値群それぞれの分散である．

Cliff & Ord は，次に述べる手続きによって，空間的自己相関を有する2変数を実験的に生成し，その成果に基づく t 確率限界値を求め，独立的な変数の場合の同様の限界値（これは t 分布表に記されている）と比較している．

ステップ1：3×3，5×5，7×7，10×10の4種類の方格メッシュの人工的な地区およびアイルランド（ダブリン郡は除く）の郡単位の地区[2]を取りあげ，それぞれにおいて2変数 X_1 と X_2 のサンプル値が計測されるとする．前者の人工的な4地区については，境界効果を除くためトーラス面に写像する．接合係数に関して，メッシュ方格または郡 i と j が辺または境界を共有するときは $w_{ij}=1$，そうでないときは $w_{ij}=0$ とし，$\sum_i w_{ij}=1$ となるように尺度化する（後述する w_{ij} はすべて尺度化されたものであり，$i \neq j$ である）．

ステップ2：X_1 と X_2 それぞれに対して，下記のような10対（つい）の自己相関パラメータ (ρ_1, ρ_2) を導入する．すなわち，(0.0, 0.0)，(0.0, 0.5)，(0.0, 0.9)，(0.5, 0.5)，(0.5, 0.9)，(0.9, 0.9)，(−0.5, −0.5)，(−0.9, −0.9)，(0.5, −0.5)，(0.9, −0.9) の10対であり，各パラメータ値は X_1 と X_2 それぞれにおける自己相関の水準を示す[3]．

第1表　独立的変数と空間的自己相関変数に関する t 確率限界値（7×7方格メッシュの場合）

パーセント点 α	独立的変数の t 分布	空間的自己相関変数の t 分布									
		(0.0, 0.0)	(0.0, 0.5)	(0.0, 0.9)	(0.5, 0.5)	(0.5, 0.9)	(0.9, 0.9)	(−0.5, −0.5)	(−0.9, −0.9)	(0.5, −0.5)	(0.9, −0.9)
0.10	1.29	1.30	1.88	7.30	2.39	7.21	8.51	0.77	0.37	1.77	5.50
0.05	1.66	1.61	2.53	9.20	3.01	9.08	10.98	0.99	0.47	2.25	7.08
0.025	1.98	1.97	2.94	10.90	3.63	11.05	12.88	1.17	0.56	2.65	8.12
0.010	2.36	2.23	3.40	13.22	4.05	13.38	14.99	1.35	0.65	3.14	10.02
0.005	2.62	2.42	3.71	14.69	4.47	16.63	16.35	1.47	0.73	3.53	11.48

注：カッコ内数値は2変数それぞれの自己相関パラメータ ρ_1 と ρ_2.　(Cliff & Ord, 1975a より)

ステップ3：上の2つのステップにおいて与えられた w_{ij} および $\rho_k(k=1, 2)$ に基づいて，次式で定義される自己相関オペレータを求める.

$$(\mathbf{I}-\rho_k\mathbf{W}_k)^{-1} \tag{3}$$

ここでの $\mathbf{I}$ は単位行列，$\mathbf{W}$ は w_{ij} を要素とする $n\times n$ の接合係数行列である.

ステップ4：標準正規母集団からの乱数として攪乱項ベクトル $\mathbf{e}_k$ を生成する.

ステップ5：相関オペレータと攪乱項ベクトルから，空間的自己相関を有する変数値ベクトル $\mathbf{x}_k$ を次式によって導く.

$$\mathbf{x}_k=(\mathbf{I}-\rho_k\mathbf{W}_k)^{-1}\mathbf{e}_k \tag{4}$$

ステップ6：ステップ4と5の作業を1,200回反復する.

このような手続きによってシミュレートされた，自己相関を有する変数値に関する t 確率限界値と，独立的な変数値に関するそれとを対比すると，第1表のようになる（7×7方格メッシュの場合）. この表から明らかなように，両種の変数間の t 値の差は著しく，空間的自己相関を有する変数に対する t 検定は，誤った推測を導くことになろう. ここで注目すべきは，X_1 と X_2 の両方または一方にプラスの自己相関がみられる場合は，非相関の場合に比べて t 値が大となり，相関が高水準になるほど t 値が増大すること，および，2変数の両方にマイナスの自己相関がある場合は，逆に t 値は小となり，相関が高水準になるほど減少することである. このことは，プラスの自己相関がある場合は，帰無仮説の棄却域が拡大し，式(1)と(2)に拠るならば，帰無仮説が棄却される確率が増大し，マイナスの自己相関がある場合は，正反対の状況が生ずることを意味する.

このような空間的自己相関の t 統計量に対する偏性が生ずる原因は，その相関を有する変数の平均が独立的変数のそれと異なることに求められる. すなわち，ある地区の変数値 x_i が，直接隣合った地区のそれ x_j と相関するという1階の自己相関がみられ，その相関がどの方向についても同一であるとすれば，この自己相関は次のように表わせる.

$$x_i=\rho\sum_{j=1}^{n}w_{ij}x_j+e_i \quad (i=1, 2, \cdots, n) \tag{5}$$

ここでの ρ, w_{ij}, e_i は以前に定義されたものと同様である. このような変数 X の平均は，最尤推定法によると，ρ が既知の場合は次のように推定される.

$$\hat{\mu}=\frac{\mathbf{x}'\mathbf{A1}}{\mathbf{1}'\mathbf{A1}}=\frac{\sum_{i=1}^{n} a_{i.}\,x_i}{\sum_{i=1}^{n} a_{i.}} \tag{6}$$

ここでの $\mathbf{A}=(\mathbf{I}-\rho\mathbf{W})'(\mathbf{I}-\rho\mathbf{W})$, $\mathbf{1}$ は 1 を要素とする列ベクトル, a は $\mathbf{A}$ の要素, $a_{i.}=\sum_{j=1}^{n} a_{ij}$ である. また, X の分散の最尤推定量 $\hat{\sigma^2}$ は,

$$\hat{\sigma^2}=n^{-1}\sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}(x_i-\hat{\mu})(x_j-\hat{\mu}) \tag{7}$$

である. これらの平均と分散は, 通常のものと比べて a_{ij} によって加重されていることが理解できるであろう.

2. 回帰パラメータについて

空間的自己相関の回帰パラメータに対する影響は, 回帰モデルを構成する変数や誤差項に何らかの系列的な自己相関がみられる場合に, 通常の最小2乗法によって回帰パラメータを推定すると, その推定値は最良不偏性をもたなくなる, ということであり, 古くから Student (1914) や Yule (1926) などによって指摘されてきた問題である. とりわけ, 時系列事象を扱う計量経済学では最重要問題の1つとされ, 多数の成果が報告されている[4]. 空間系列の自己相関に関する研究は, Cliff & Ord (1972) 以来活発になってきている. このような自己相関の回帰パラメータへの影響に対する注目の増大は, 回帰分析が諸分野できわめて頻用されていることに由来しているのはいうまでもない.

線形回帰モデルは, Y を従属変数, m 個の X を独立変数, β を未知の回帰パラメータ, e を誤差, i を観測体とすると, 一般に次のように表現される.

$$Y_i=\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}+\cdots+\beta_m x_{mi}+e_i\ (i=1, 2, \cdots, n) \tag{8}$$

または, $\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\mathbf{e}$ (8′)

しかし, これには4つの基本的な仮定が設けられている. それは, (1) $\mathrm{E}(\mathbf{e})=0$, (2) $\mathrm{E}(\mathbf{ee}')=\sigma^2\mathbf{I}$, (3) $\mathbf{X}$ は固定された数の集合であること, (4) $\mathbf{X}$ の階数 $<n$ であること, である. (1)は, e_i の期待値がすべてゼロであることを述べており, 不偏性の仮定といわれる. (2) は, e_i がすべて一定の分散 σ^2 をもつこと, および e_i を2つずつ取りあげた場合に互いに独立的であることを意味しており, 分散一様性および独立性の仮定といわれる. (3) は, $\mathbf{Y}$ の変動が $\mathbf{e}$ の変動によってもたらされるものであり, $\mathbf{X}$ は $\mathbf{e}$ とは独立していることを述べている. (4) は, サンプル (観測体) の個数が推定されるべき回帰パラメータの個数より大であり, かつ m 個の X の間は無相関であることを述べている. 回帰モデルにおける未知の $\boldsymbol{\beta}$ の推定法として, 最小2乗法が最も著名であり, どの統計学の教科書にもこの推定法が説明されている. この方法は, 上述の4つの仮定のもとで, e_i をすべてのサンプルについて最小にするような推定量 $\hat{\boldsymbol{\beta}}$ を求める方法である. すなわち, $L=\sum_{i=1}^{n} e_i^2$ とおけば, L の行列形式は次のように記すことができる.

$$L=\mathbf{e}'\mathbf{e}=(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta})'(\mathbf{Y}-\mathbf{X}\boldsymbol{\beta}) \tag{9}$$

これは次のように改書することができる.

$$L = \mathbf{Y}'\mathbf{Y} = 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \tag{10}$$

この式のLを$\boldsymbol{\beta}$に関して偏微分して得られる次の偏導関数をゼロとおき，

$$\frac{\partial L}{\partial \boldsymbol{\beta}} = -2\mathbf{X}'\mathbf{Y} + 2\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \tag{11}$$

それを解けば，次式が導かれる．

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{Y} \tag{12}$$

上式から得られる$\widehat{\boldsymbol{\beta}}$は最良線形不偏推定量である．「最良」とは，この推定量がすべての線形不偏推定量のうち最小の分散をもっていることを意味し，推定値としては最も正確なものである．また，「線形」とは，推定量が$\mathbf{Y}$の測定値の線形結合（１次関数）として表現できることを意味する[5)]．最良という性質は推定量としては最も望まれるものであり（正確な推定値はその典型である），線形という性質は上述の(3)の仮定を，不偏という性質は(1)の仮定をそれぞれ反映している．また，$\widehat{\boldsymbol{\beta}}$が得られることは，式(12)にみられる逆行列$(\mathbf{X}'\mathbf{X})^{-1}$が存在することに他ならないので，この推定量は(4)の仮定をも満たすことになる．このようなことから，最小２乗法は回帰パラメータの推定法としてきわめて好ましいものとの評価を受けているのである．

しかし，空間的自己相関を示す事象に関する回帰モデルに対して最小２乗法を利用した場合はどうなるであろうか．地理的事象は，前報告で述べたように，空間的自己相関を有する場合が多いのである．従属変数$\mathbf{Y}$で表わされる事象に自己相関がみられるとすると，上述の仮定(3)から，その相関は誤差項$\mathbf{e}$にみられることになる．そうすると，仮定(2)を満たしえなくなることが容易に理解できるであろう．つまり，回帰パラメータの最良線形不偏推定量が得られなくなるわけである．それでは，どのような推定量が得られるあろうか．式(8)の回帰モデルを次のようであるとしよう．

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \tag{13}$$

そして，次式で定義されるような１階の自己相関が$\mathbf{u}$にみられるとする．

$$u_i = \rho \sum_{j=1}^{n} w_{ij}\, u_j + e_i \tag{14}$$

$$\text{または，}\ \mathbf{u} = \rho\mathbf{W}\mathbf{u} + \mathbf{e} \tag{14$'$}$$

上式での$\mathbf{e}$については，$\mathrm{N}(\mathbf{O}, \sigma^2\mathbf{I})$の分布に従うものとする．$\rho$と$\mathbf{W}$は以前の定義と同様である．式(14′)は次のように改書することができる．

$$\mathbf{u} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{e} \tag{15}$$

この式での行列$(\mathbf{I} - \rho\mathbf{W})$は正値定符号であるとすれば，$\mathbf{u}$の期待値と分散は次のようになる．

$$\mathrm{E}(\mathbf{u}) = 0 \tag{16}$$

$$\mathrm{E}(\mathbf{u}\mathbf{u}') = \sigma_u^2\,[(\mathbf{I} - \rho\mathbf{W})(\mathbf{I} - \rho\mathbf{W})']^{-1} \tag{17}$$

簡単にするために，上式での$[(\mathbf{I} - \rho\mathbf{W})(\mathbf{I} - \rho\mathbf{W})']^{-1}$を$\mathbf{M}$とおくと，

$$\mathrm{E}(\mathbf{u}\mathbf{u}') = \sigma_u^2\mathbf{M} \tag{18}$$

明らかに$\mathbf{M} \neq \mathbf{I}$であるので，回帰モデルの誤差項は，非スカラーの共分散行列をもつことになる．このことは，仮定(2)を乱すことを意味するであろう．式(13)に最小２乗法を適用すると，$\widehat{\boldsymbol{\beta}}$は次式で与えられる．

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} \tag{19}$$

このことから，$\widehat{\beta}$の期待値は次のようになる．

$$\mathrm{E}(\widehat{\beta})=\mathrm{E}[(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\boldsymbol{\beta}]+\mathrm{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}]$$
$$=\boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{E}(\mathbf{u}) \tag{20}$$

$\mathrm{E}(\mathbf{u})=0$ であるので，$\mathrm{E}(\widehat{\beta})=\boldsymbol{\beta}$ である．したがって，$\mathbf{X}$と$\mathbf{u}$が互いに独立的，つまり $\mathrm{E}(\mathbf{X}'\mathbf{u})=0$ であるならば，$\widehat{\beta}$は，自己相関が存在していても不偏推定量である．$\widehat{\beta}$の分散については，式（19）から次のようになる[6]．

$$\mathrm{E}(\widehat{\beta}-\boldsymbol{\beta})(\widehat{\beta}-\boldsymbol{\beta})'=\mathrm{E}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}]$$
$$=\sigma_u^2(\mathbf{X}'\mathbf{X})^{-1}[(\mathbf{X}'\mathbf{M}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}] \tag{21}$$

この式での最初の項$\sigma_u^2(\mathbf{X}'\mathbf{X})^{-1}$は，自己相関がみられないときの$\widehat{\beta}$の分散である．$[(\mathbf{X}'\mathbf{M}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}]$を調べると，この項は，$\rho=0$ の場合，つまり誤差項に自己相関がみられない場合，および独立変数に自己相関がない場合には$\mathbf{I}$になることがわかる．したがって，誤差項に自己相関がみられるか否かは別として独立変数に自己相関がなければ，最小2乗法による回帰パラメータの推定量は最良推定量になるといえる．しかし，このような状況は実際にはほとんどなく，もし，$\rho>0$であり，1つまたはそれ以上の独立変数が自己相関を有するならば，上述の行列の各要素は1より大となり，最小2乗法による推定量は不偏性をもつものの，その分散は真の分散に比べて過小に推定されることになる（このことは後述の数値例によって確かめられる）．要するに，$(\mathbf{X}'\mathbf{M}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}=\mathbf{I}$ の場合以外には，最小2乗推定量は，$[(\mathbf{X}'\mathbf{M}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}-\mathbf{I}]$ の量だけ偏るわけである．

このような$\widehat{\beta}$の分散に対する偏りに加えて，いま1つの偏りが生ずる．それは，誤差項$\mathbf{u}$の分散に対してのものである．この分散の最小2乗推定量 S^2 は，回帰からの最小2乗推定誤差のベクトルを$\widehat{\mathbf{u}}$とすると，次式で表わせる．

$$S^2=\frac{\widehat{\mathbf{u}}'\widehat{\mathbf{u}}}{n-m-1} \tag{22}$$

これの期待値は次のようである．

$$\mathrm{E}(S^2)=\frac{1}{n-m-1}\mathrm{E}(\widehat{\mathbf{u}}'\widehat{\mathbf{u}})$$
$$=\frac{\sigma_u^2}{n-m-1}\mathrm{E}\{\mathrm{tr}\mathbf{M}-\mathrm{tr}[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{M}\mathbf{X}]\} \tag{23}$$

このことから，誤差項のプラスの自己相関がある場合は，大カッコの表現式の$(n-m-1)$からの差だけ，誤差の分散は過小に推定されるといえる．この種の推定量は，回帰の有意性を検定する際に用いられるt統計量やF統計量の分母に現われるので，誤差の分散の過小推定は，t値やF値を過大にする．また，重相関係数の定義式にも誤差の分散が記されているので，この係数をも過大にするのである．

Ⅲ　空間的自己相関の除去

上述のような既存の各種の統計量に対する空間的自己相関の影響は，既存の統計量に頼る研究者にとっては明らかにノイズといえる．したがって，このノイズを取り除く方法が当然のことながら要請されるであろう．しかし，このことは，投入データ値に空間的自己相関が存在するか否かの見定めの

後のことである．この見定めは，前報告で述べられた空間的自己相関に関する諸測度の適用によって行なわれるが，それに関して注目すべき事柄が2つある．1つは，Cliff & Ord (1975b) や Haining (1977, 1978) などの研究成果をもとにして，空間的自己相関に関する諸測度の検定力について Cliff & Ord が下した次のような結論である (Cliff & Ord, 1981, p. 178). すなわち，(1)データ数値が区間尺度の場合は，Moran の I 統計量が，Geary の C 統計量や BB, BW のような接合統計量より秀れている．このことは，正規分布のデータについてのみ成立するが，データのタイプより単位地区の形状の方がむしろ重要である．(2)順序尺度の場合は，I 統計量を利用すべきである．(3)バイナリー尺度の場合は，BW 統計量を利用すべきである．しかし，黒色 (B) をもつ地区が全体のほぼ50%を占める状況においての利用が望ましい，というのがそれである．いま1つは，Brandsma & Ketellapper (1979) が指摘しているように，回帰モデルの誤差（残差）項に関する空間的自己相関の有無の検討には，次の統計量を用いることが望ましいということである[7].

$$I_k=\frac{\hat{\mathbf{e}}\,\mathbf{W}\,\hat{\mathbf{e}}}{\hat{\mathbf{e}}'\hat{\mathbf{e}}} \tag{24}$$

このような統計量の適用によってデータに自己相関の存在が認められた場合，データからこの相関を除去することが要請されるが，それには2つの戦術が考えられる．1つは，自己相関が存在するデータ数値を無相関のそれに変換する方法であり，いま1つは，既存の統計量を，自己相関が存在していても正しい推測結果を導く統計量に修正する方法である．前者に関しては減算法 (differencing method), 後者に関しては，前章第1節に対応する修正 t 統計量，および前章第2節に対応する一般化最小2乗法と最尤法をそれぞれ述べる．

1. 減　算　法

この方法は，基本的には，データ数値 x_i(i=1, 2, …, n) から空間的自己相関成分を減ずることによって，n 個のサンプルについて互いに独立した数値に変換しようとする方法である．すなわち，空間的自己相関成分は $\rho\sum_{j=1}^{n}w_{ij}\,x_j$ によって表現されると仮定して，

$$\Delta x_i=x_i-\rho\sum_{j=1}^{n}w_{ij}\,x_j \tag{25}$$

または，$\mathbf{X}^*=\mathbf{X}-\rho\mathbf{W}\mathbf{X}=(\mathbf{I}-\rho\mathbf{W})\,\mathbf{X}$ (25′)

なる1階の階差式によって x_i を Δx_i に変換するのである．一般的には，$\rho=1$ とおいたときの次式が用いられるとされている (Upton & Fingleton, 1985, p. 283).

$$\mathbf{X}^*=\mathbf{X}-\mathbf{W}\mathbf{X}=(\mathbf{I}-\mathbf{W})\,\mathbf{X} \tag{26}$$

しかし，ここでは $0<|p|\leqslant 1$ として論をすすめる．この式を調べると，Δx_i は，サンプル地区 i と隣接したすべての地区 j での x_i の加重平均を x_i から減じたものであることが理解されよう．

この減算法は，データ数値の独立化に対してどのような効果があるであろうか．Martin (1974) は，回帰パラメータの推定問題と関連させて，この効果を実験的に評価している．すなわち，初めに回帰モデルを次のように想定しておき，

$$y_i=1.0+0.5\,x_i+u_i \qquad (i=1,\,2,\,\cdots,\,n\ \text{方格}) \tag{27}$$

x_i と u_i それぞれが次式で示されるような空間的自己相関を示すとする．

$$x_i=\lambda \sum_j w_{ij}\, x_j+v_i \tag{28}$$

$$\text{または，}\ \mathbf{X}=\lambda\mathbf{W}\mathbf{X}+\mathbf{v} \tag{28'}$$

$$u_i=\rho \sum_j w_{ij}\, u_j+e_i \tag{29}$$

$$\text{または，}\ \mathbf{u}=\rho\mathbf{W}\mathbf{u}+\mathbf{e} \tag{29'}$$

これら2式について，$\mathbf{v}$ の分布は $\mathrm{N}(\mathbf{O}, \sigma_v^2\mathbf{I})$, $\mathbf{e}$ の分布は $\mathrm{N}(\mathbf{O}, \sigma_e^2\mathbf{I})$ であると仮定する．次いで，下記のような手続きによって $\mathbf{X}$ と $\mathbf{u}$ をシミュレートさせる．当然のことながら，$\mathbf{X}$ と $\mathbf{u}$ が得られると，$\mathbf{Y}$ もまた式(27)から得られる．

ステップ1：サンプル地区として，5×5，6×6，7×7の3種の方格メッシュを取りあげる．

ステップ2：式(28)の λ として 0.0，0.4，0.8，式(29)の ρ として0.0，0.2，0.4，0.6，0.8，1.0 を与える．

ステップ3：正規分布すると仮定された $\mathbf{v}$ について平均75，分散201，$\mathbf{e}$ については平均0，分散251 をそれぞれとるように，$\mathbf{v}$ と $\mathbf{e}$ をサンプル地区数だけ乱数を発生させる．

ステップ4：$\mathbf{X}$ を次式から計算する．

$$\mathbf{X}=(\mathbf{I}-\lambda\mathbf{W})^{-1}\mathbf{v} \tag{30}$$

これによって，$|\lambda|>0$ の場合は $\mathbf{X}$ は空間的自己相関を有することになる．同様に $\mathbf{u}$ を次式から計算する．

$$\mathbf{u}=(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{e} \tag{31}$$

ステップ5：計算された $\mathbf{X}$ と $\mathbf{u}$ を式(27)に代入して $\mathbf{Y}$ を求める．

ステップ6：このような手続きによってシミュレートされた $\mathbf{Y}$ と $\mathbf{X}$ の数値に対して最小2乗法を適用し，$\hat{\beta}$ とそれの分散 $\mathrm{var}(\hat{\beta})$ を求める．

ステップ7：同様の数値を減算法によって独立的数値に変換し，その結果に対して最小2乗法を適用し，$\hat{\beta}$ と $\mathrm{var}(\hat{\beta})$ を求める．

ステップ8：ステップ3～7を100回反復し，変換前の場合の $\hat{\beta}$ と $\mathrm{var}(\hat{\beta})$ それぞれの平均値と，変換後の場合のそれらを求める．そして，それらの結果値を両方の場合について比較する．

この比較の結果は第2表と第3表に示されるとおりである．両表にみられる平均平方誤差（MSE）は，$\mathrm{E}(\hat{\beta}-\text{真の }\beta)^2$ と定義されるものであり（真の β は式(27)にみられる0.5である），$\hat{\beta}$ の真の分散を意味する．これを基準にすれば，$\mathrm{var}(\hat{\beta})$ の偏りの程度を評価することができる．

第2表は，$\mathbf{Y}$ と $\mathbf{X}$ の両方に空間的自己相関がある（したがって $\mathbf{u}$ にも同様の相関がある）場合の $\hat{\beta}$, $\mathrm{var}(\hat{\beta})$, MSE を示している．この表によると，$\hat{\beta}$ に関しては，λ と ρ が大になるに伴い，真の β 値からの偏りは全体的に拡大し，その傾向は地区数が小なほど著しいこと，しかし，その偏りが真の β 値を上回るものか否かについては一定の傾向が認められないことが読み取れる．$\mathrm{var}(\hat{\beta})$ に関しては興味ある4つの傾向がみられる．それは，(1) $\mathrm{var}(\hat{\beta})$ は真の分散に比べて全体的に過小に評価されている．(2)地区数と λ を一定にすると，ρ の増大に伴って $\mathrm{var}(\hat{\beta})$ は大になり，真の分散からの偏りは拡大する．ρ が0.6以上になるとその偏りは急激に拡大し，(5×5) の場合では $\rho=1.0$ のと

第２表　非変換の場合の最小２乗推定結果

空間的自己相関パラメータ										
λ	ρ	(5×5)			(6×6)			(7×7)		
		$\hat{\beta}$	var $(\hat{\beta})$	MSE	$\hat{\beta}$	var $(\hat{\beta})$	MSE	$\hat{\beta}$	var $(\hat{\beta})$	MSE
0.0	0.0	0.515	0.127	0.136	0.495	0.085	0.091	0.524	0.061	0.065
	0.2	0.577	0.126	0.139	0.574	0.086	0.085	0.493	0.066	0.069
	0.4	0.534	0.131	0.145	0.466	0.101	0.107	0.525	0.069	0.070
	0.6	0.436	0.190	0.217	0.527	0.111	0.118	0.541	0.084	0.099
	0.8	0.514	0.246	0.381	0.544	0.176	0.178	0.499	0.138	0.154
	1.0	0.631	0.504	1.197	0.488	0.627	0.914	0.561	0.608	0.877
0.4	0.0	0.534	0.113	0.124	0.495	0.074	0.075	0.534	0.052	0.059
	0.2	0.481	0.113	0.153	0.490	0.077	0.097	0.497	0.058	0.059
	0.4	0.479	0.119	0.159	0.473	0.089	0.103	0.529	0.060	0.063
	0.6	0.488	0.154	0.207	0.531	0.098	0.115	0.533	0.073	0.115
	0.8	0.433	0.251	0.552	0.529	0.153	0.233	0.503	0.136	0.193
	1.0	0.481	0.635	1.264	0.490	0.501	0.880	0.514	0.494	0.861
0.8	0.0	0.519	0.071	0.080	0.492	0.042	0.044	0.516	0.027	0.034
	0.2	0.538	0.077	0.099	0.528	0.043	0.055	0.499	0.030	0.038
	0.4	0.568	0.079	0.124	0.490	0.048	0.078	0.513	0.032	0.044
	0.6	0.502	0.088	0.140	0.531	0.053	0.091	0.516	0.038	0.090
	0.8	0.495	0.153	0.402	0.516	0.082	0.215	0.503	0.071	0.211
	1.0	0.476	0.454	0.965	0.513	0.407	0.800	0.490	0.349	0.736

(Martin, 1974 より)

き var$(\hat{\beta})$ は真の分散の 1/2 にも減じている．（３）地区数と ρ を一定にすると，λ の増大につれて var$(\hat{\beta})$ は ρ の場合とは逆に減少する．しかし，真の分散からの偏りは拡大する．例えば，(6×6) での $\rho=1.0$ の場合，$\lambda=0.0$ のときの var$(\hat{\beta})$ は真の分散の 7/10 であるが，$\lambda=0.8$ のときでは 1/2 にも減少している．(4) ρ と λ を一定にすると，地区数が大になるほど var$(\hat{\beta})$ は小になり，これの偏りもまた縮小する．これらのことから，一定の地区数のもとでは，誤差項での自己相関は，その相関の増大に伴って var$(\hat{\beta})$ を大にさせ，他方，独立変数での自己相関は，その相関の増大に伴ってその分散を逆に小さくさせるという正反対の効果を示すが，いずれの自己相関も var$(\hat{\beta})$ を真の分散より小にさせ，相関が高くなるにつれて両者の差を拡大させるといえる．

第３表に関しては，$\hat{\beta}$ の偏りは第２表の場合より縮小し，それは地区数がふえるにつれて著しくなる．このことは，減算法が一層正確な $\hat{\beta}$ を導いていることを示唆している．var$(\hat{\beta})$ については，第２表の場合と若干異なる傾向を呈している．それは，(1) 一定の λ のもとで，var$(\hat{\beta})$ の値とその偏りはともに ρ の増大に伴って減少し，$\rho \geqq 0.6$ のときに偏りがきわめて小になっている．(2) var$(\hat{\beta})$ の範囲が全体的に縮小している．このことは，減算法によって $\hat{\beta}$ が最良性を示すようになることを物語る．(3) λ の増大に伴う var$(\hat{\beta})$ の値の変動は第２表の場合とは異なり，増加傾向を示すが，その偏りは ρ の場合と同様に縮小している．これらのことから，減算法は偏りを相当改善させているとい

第 3 表　変換の場合の最小 2 乗推定結果

空間的自己相関パラメータ										
λ	ρ	(5×5)			(6×6)			(7×7)		
		$\hat{\beta}$	var $(\hat{\beta})$	MSE	$\hat{\beta}$	var $(\hat{\beta})$	MSE	$\hat{\beta}$	var $(\hat{\beta})$	MSE
0.0	0.0	0.547	0.132	0.147	0.529	0.090	0.146	0.489	0.062	0.099
	0.2	0.559	0.107	0.145	0.509	0.072	0.118	0.513	0.054	0.088
	0.4	0.510	0.106	0.144	0.451	0.068	0.096	0.501	0.047	0.068
	0.6	0.450	0.100	0.130	0.517	0.062	0.069	0.541	0.044	0.057
	0.8	0.470	0.085	0.086	0.507	0.056	0.060	0.490	0.043	0.056
	1.0	0.489	0.082	0.082	0.497	0.055	0.056	0.500	0.050	0.056
0.4	0.0	0.555	0.172	0.236	0.507	0.118	0.162	0.529	0.082	0.109
	0.2	0.507	0.138	0.179	0.497	0.099	0.135	0.486	0.071	0.103
	0.4	0.479	0.117	0.145	0.529	0.094	0.124	0.505	0.061	0.080
	0.6	0.516	0.116	0.140	0.486	0.086	0.111	0.556	0.057	0.074
	0.8	0.496	0.104	0.138	0.522	0.077	0.086	0.490	0.056	0.063
	1.0	0.506	0.104	0.103	0.511	0.076	0.078	0.515	0.056	0.056
0.8	0.0	0.495	0.185	0.191	0.521	0.129	0.173	0.558	0.090	0.099
	0.2	0.536	0.157	0.163	0.507	0.104	0.127	0.498	0.078	0.101
	0.4	0.532	0.142	0.143	0.496	0.098	0.112	0.510	0.068	0.076
	0.6	0.482	0.130	0.131	0.520	0.092	0.103	0.516	0.038	0.070
	0.8	0.481	0.127	0.127	0.513	0.086	0.096	0.489	0.034	0.068
	1.0	0.500	0.124	0.123	0.500	0.084	0.086	0.498	0.032	0.033

(Martin, 1974より)

えるであろう．

減算法の実際問題に対する利用に関して，Cliff & Ord は，前述の平均間の差の有意性検定の議論および上述の Martin による成果を踏えて，次のような限定的ではあるが，有用な指針を提示している (Cliff & Ord, 1981, p. 193)．すなわち，(1) マイナスの自己相関がみられる場合は，その相関を含むデータ数値を利用すると，帰無仮説を棄却する可能性が小さくなる．(2) 無相関に近い場合は，自己相関を含むデータ数値の利用は，帰無仮説を棄却する可能性をわずかに高め，減算法によって変換されたデータ数値は，その可能性を若干減少させる．(3) プラスの高自己相関を示す場合は，被変換のデータ数値の利用は，帰無仮説を棄却する可能性を若干低下させ，自己相関を示すデータ数値を利用する場合より検定は信頼の高いものとなる．

このことは，マイナスの自己相関がみられる場合は，その相関を含むデータ数値を利用しても推測結果に偏りが生ずることはないが，プラスの自己相関がみられる場合は，減算法による変換を施すべきであることを示唆している．Martin の研究結果から明らかなように，とりわけ $0.4 < \rho \leqq 1.0$ の場合は減算法は有効である．しかし，注意すべきは，1 つの平均に関する有意性検定に対してはこの方法が利用できないことである．それは，$\sum_{j=1}^{n} w_{ij} = 1$ とされているため，Δx_i の平均はゼロとなるからである．

減算法の利用に際して1つの問題となることは，データ数値にみられる空間的自己相関の方向と強度を表わす係数ρをどのように見定めるかである．この問題は後述する事項にも深く関連する．係数ρの最も正式な推定法は最尤法であるが，これの難点は煩瑣な手続きを伴うことである．そこで，Cliff & Ord (1981)は簡便な推定法を提唱している．それはMoranのI統計量を利用する方法である．しかし，この統計量は，その定義式から理解されるように[8]，ピアソン積率相関係数の場合のような［−1，＋1］の範囲を取らず，この範囲を逸脱することがしばしば起こる．それゆえ，I統計量を直截的に$\hat{\rho}$とすることはできない．それに対処するために，Cliff & Ordは次のような変換式を用意している．

$$\hat{\rho}=\frac{I}{\max|I|} \tag{32}$$

ここでの$\max|I|$は次式によって定義されるものである．

$$\max|I|=\left[\frac{\mathrm{var}\left(\sum_{j=1}^{n} w_{ij} z_j\right)}{\mathrm{var}(z_i)}\right]^{1/2} \qquad (i \neq j) \tag{33}$$

ただし，$z_i=x_i-\bar{x}$である．このような推定法は実用上きわめて便利なものであるが，Cliff & Ord (1975a)が指摘しているように，この方法による$\hat{\rho}$はあくまでも近似値であり，サンプル地区数が25以上の状況にその使用が限定されることに注意すべきである．

2.　修正t統計量

前章で述べた平均間の差の有意性検定に関するt統計量に対する空間的自己相関のノイズ効果を除くために，Cliff & Ord (1975a)は修正t統計量というべきものを考案している．この統計量は，自己相関を含むデータ数値に適用しても偏りを生じないものである．彼らは初めに，$N(\mu_k, \sigma^2 \mathbf{A}_k^{-1})$の分布をもつ変数$X_k$ ($k=1, 2, \cdots, m$)を想定する．ここでの$\mathbf{A}_k=(\mathbf{I}-\rho_k\mathbf{W}_k)'(\mathbf{I}-\rho_k\mathbf{W}_k)$である．つまり，$m$個の変数$X$はそれぞれ異なる方向と強度の1階の空間的自己相関を呈するとするのである．そうすると，前掲の式(6)と(7)の拡張として，この変数の平均と分散の推定値が次のように定義できる．

$$\hat{\mu}_k=\frac{\mathbf{x}'_k\mathbf{A}_k\mathbf{1}}{\mathbf{1}'\mathbf{A}_k\mathbf{1}} \tag{34}$$

$$\begin{aligned}\hat{\sigma}^2=(n_1+n_2+\cdots+n_m)^{-1}\Big[&\sum_j\sum_i a_{ij}^{(1)}(x_{1i}-\hat{\mu}_1)(x_{1j}-\hat{\mu}_1)\\ &+\sum_j\sum_i a_{ij}^{(2)}(x_{2i}-\hat{\mu}_2)(x_{2j}-\hat{\mu}_2)+\cdots\cdots\\ &+\sum_j\sum_i a_{ij}^{(m)}(x_{mi}-\hat{\mu}_m)(x_{mj}-\hat{\mu}_m)\Big]\end{aligned} \tag{35}$$

次いで，これに基づいて，2変数($k=2$)の場合に関する母平均間の差には有意性はないという帰無仮説H_0: $\mu_1=\mu_2$の対立仮説H_1: $\mu_1 \neq \mu_2$に対する検定を考えると（いずれの仮説でもσ^2は未知），尤度比の利用から次式が導かれる．

$$t=\frac{\widehat{\mu}_1-\widehat{\mu}_2}{\widetilde{\sigma}\,(m_1^{-1}+m_2^{-1})^{1/2}} \tag{36}$$

ここでの分散（不偏）推定量 $\widetilde{\sigma^2}=(n_1+n_2)\,\widehat{\sigma^2}/(n_1+n_2-2)$，$m_k=\mathbf{1}'\mathbf{A}_k\mathbf{1}=\sum_j\sum_i a_{ij}^{(k)}$ である．利用しやすくするために，$\widehat{\mu}_k=\bar{x}_k$，$m_k=n_k(1-\widehat{\rho}_k)^2$ として上式を書き改めると，次のようになる．

$$t=\frac{\bar{x}_1-\bar{x}_2}{\widetilde{\sigma}\left(\dfrac{1}{n_1(1-\widehat{\rho}_1)^2}+\dfrac{1}{n_2(1-\widehat{\rho}_2)^2}\right)^{1/2}} \tag{37}$$

$$\widetilde{\sigma^2}=(n_1+n_2-2)^{-1}\,[\mathbf{x}_1{}'\mathbf{A}_1\mathbf{x}_1+\mathbf{x}_2{}'\mathbf{A}_2\mathbf{x}_2-\bar{x}_1^2(1-\widehat{\rho}_1)^2-\bar{x}_2^2(1-\widehat{\rho}_2)^2] \tag{38}$$

これが修正 t 統計量である．

この統計量の有効性を検討するために，Cliff & Ord は，前述の手続きに従って生成された空間的自己相関を有する2変数の数値を式(37)と(38)に投入し，5つの有意水準ごとの t 確率限界値を求めている．さらに，分散推定量について次式で定義される通常のものを用意し，

$$\widetilde{\sigma^2}=\frac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2} \tag{39}$$

変数値を式(37)と(39)にも投入して同様の限界値を求めている（式(39)の記号は前掲の式(2)と同様である）．7×7方格メッシュの場合での数値を用いたときの結果は，第4表のとおりである．なお，この作業に際しての ρ_1 と ρ_2 の推定式は前節で述べられた方法によるものである．第4表を調べると，自己相関がみられるデータ数値を通常の t 統計量に投入して場合（第1表参照）と比べて，偏りは相当除去され，修正 t 統計量による確率限界値が通常のそれに接近していることが読み取れる．しかし，ρ_1 と ρ_2 がともにあるいは一方がプラスの場合は，式(38)を用いたときの限界値（表中の欄a）は依然としてかなりの偏りを示し，式(39)を用いたときの方が改善の程度がはるかに高い．ρ_1 と ρ_2 がと

第4表　修正 t 統計量に関する t 確率限界値（7×7方格メッシュの場合）

パーセント点 α	独立的変数の t 分布	修正 t 統計量に関する t 分布 (0.0, 0.0) a	b	(0.0, 0.5) a	b	(0.0, 0.9) a	b	(0.5, 0.5) a	b	(0.5, 0.9) a	b
0.10	1.29	1.36	1.34	1.41	1.28	2.35	1.27	1.51	1.34	2.33	1.26
0.05	1.66	1.72	1.69	1.96	1.86	4.04	1.59	1.93	1.74	3.97	1.65
0.025	1.98	1.96	1.94	2.34	2.17	6.08	2.12	2.34	2.09	6.34	2.14
0.01	2.36	2.34	2.35	2.76	2.61	10.09	3.28	2.71	2.49	10.12	2.60
0.005	2.62	2.56	2.54	2.85	2.70	20.34	3.55	2.98	2.67	11.00	2.87
α	t	(0.9, 0.9) a	b	(−0.5, −0.5) a	b	(−0.9, −0.9) a	b	(0.5, −0.5) a	b	(0.9, −0.9) a	b
0.10	1.29	2.50	1.01	1.31	1.16	1.27	0.68	1.60	1.36	2.27	1.08
0.05	1.66	4.09	1.25	1.66	1.41	1.52	0.88	1.98	1.78	3.77	1.49
0.025	1.98	6.34	1.45	1.88	1.76	1.87	1.02	2.56	2.28	5.46	1.80
0.01	2.36	9.88	1.84	2.24	1.86	1.99	1.14	3.24	2.82	10.05	2.24
0.005	2.62	12.17	2.71	2.27	1.98	2.18	1.20	3.78	3.20	12.48	2.48

注：aは分散推定量として式（38）を用いた場合，bは式（39）を用いた場合である．カッコ内数値は第1表と同様．　（Cliff & Ord, 1975aより）

もにマイナスの場合は，式(38)を用いたときの方が式(39)を用いたときより偏りが少ない．このようなことから，ρ_1 と ρ_2 の両方または一方がプラスの場合は式(39)を用いた修正 t 統計量が，他方，ρ_1 と ρ_2 の両方がマイナスの場合は式(38)を用いた修正 t 統計量がそれぞれ適切であるといえる．

3.　一般化最小２乗法

線形回帰モデルにおける誤差項に空間的自己相関がみられる場合に回帰パラメータの最良線形不偏推定量を求める方法として，最も衆知のものは Aitken (1934) による一般化最小２乗法である．ここで注意すべきは，「誤差項に空間的自己相関がみられる線形回帰モデル」とは，Haggett らが述べているように，線形回帰モデルとして正しく特定化され，従属変数と独立変数それぞれのサンプル値に極値が余りみられない場合のモデルである，ということである (Haggett, *et al.*, 1977, p.p. 364～365).

一般化最小２乗法の問題は衆知のように，回帰モデルの誤差の分散共分散行列 $\mathbf{V}$ が $\mathbf{V}=\sigma^2\mathbf{D}(\mathbf{D}\neq\mathbf{I})$ の形を取るときに，回帰パラメータ $\boldsymbol{\beta}$ の最良線形不偏推定量をどのように導くかである．このことは，$\mathbf{Y}_*=\mathbf{X}_*\mathbf{b}+\mathbf{e}_*$ かつ $\mathrm{var}(\mathbf{Y}_*)=\sigma^2\mathbf{I}$ になるように，$\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\mathbf{e}$ における $\mathbf{Y}$ と $\mathbf{X}$ を，新しい変数 $\mathbf{Y}_*$ と $\mathbf{X}_*$ にそれぞれ変換することに帰着する．この変換には $\mathbf{D}$ が正値定符号であることが必要とされるが，この条件が満たされるならば，$\mathbf{D}$ は非特異行列 $\mathbf{P}$ によって次のように表現することができる．

$$\mathbf{D}=\mathbf{P}\mathbf{P}' \tag{40}$$

そこで，$\mathbf{Y}_*=\mathbf{P}^{-1}\mathbf{Y}$, $\mathbf{X}_*=\mathbf{P}^{-1}\mathbf{X}$ なる変換を行なう．そうすると，次式が得られる．

$$\mathrm{E}(\mathbf{Y}_*)=\mathbf{X}_*\mathbf{b} \tag{41}$$

$$\mathbf{e}_*=\mathbf{P}^{-1}\mathbf{e} \tag{42}$$

式(42)から次のことが判明する．

$$\begin{aligned}\mathrm{E}(\mathbf{e}_*\mathbf{e}_*')&=\mathrm{E}\,[\mathbf{P}^{-1}\mathbf{e}\,\mathbf{e}'(\mathbf{P}^{-1})']=\sigma^2\,\mathbf{P}^{-1}\mathbf{D}\,(\mathbf{P}^{-1})'\\&=\sigma^2\,\mathbf{I}\end{aligned} \tag{43}$$

かくして，$\mathbf{Y}_*=\mathbf{X}_*\mathbf{b}+\mathbf{e}_*$ に通常の最小２乗法を施せば，パラメータの最良線形不偏推定量が得られることが理解できよう．したがって，式(41)の $\mathbf{b}$ の推定量 $\widehat{\mathbf{b}}$ は次式によって与えられる（式(19)と対比されたい）．

$$\widehat{\mathbf{b}}=(\mathbf{X}_*'\mathbf{X}_*)^{-1}\mathbf{X}_*'\mathbf{Y}_* \tag{44}$$

これをもとに戻せば，次のように表わせる．

$$\widehat{\mathbf{b}}=(\mathbf{X}'\mathbf{D}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{D}^{-1}\mathbf{Y} \tag{45}$$

この推定量の分散は次のようになる．

$$\mathrm{var}\,(\widehat{\mathbf{b}})=\sigma^2(\mathbf{X}_*'\mathbf{X}_*)^{-1}=\sigma^2\,(\mathbf{X}'\mathbf{D}^{-1}\mathbf{X})^{-1} \tag{46}$$

ここでの σ^2 は次式によって推定される．

$$\widehat{\sigma^2}=\widehat{\mathbf{e}_*}'\widehat{\mathbf{e}_*}/(n-m)=\widehat{\mathbf{e}}'\mathbf{D}^{-1}\widehat{\mathbf{e}}/(n-m) \tag{47}$$

回帰モデルの誤差項に空間的自己相関がみられるという状況は，前章第２節に述べられたように，１階の自己相関を仮定すれば，次のように表わせる．

$$\mathbf{Y}=\mathbf{X}\boldsymbol{\beta}+\mathbf{u} \tag{48}$$

$$\mathbf{u}=\rho\mathbf{W}\mathbf{u}+\mathbf{e} \tag{49}$$

ただし，$\mathbf{e}$は$\mathrm{N}(\mathbf{O},\sigma^2\mathbf{I})$の分布をもつとする．式(49)は次のように改書できる．

$$\mathbf{e}=(\mathbf{I}-\rho\mathbf{W})\mathbf{u} \tag{50}$$

もし，$\mathbf{u}$の分布が$\mathrm{N}(\mathbf{O},\sigma^2\mathbf{D})$であるならば，上式から次の式が成立する（ここでの$\mathbf{D}^{-1}$は，前章第1節および本章第2節で定義された$\mathbf{A}$に他ならない）．

$$\mathbf{D}^{-1}=(\mathbf{I}-\rho\mathbf{W})'(\mathbf{I}-\rho\mathbf{W}) \tag{51}$$

$$\begin{aligned}\mathbf{V}&=\sigma^2[(\mathbf{I}-\rho\mathbf{W})'(\mathbf{I}-\rho\mathbf{W})]^{-1}\\&=\sigma^2(\mathbf{I}-\rho\mathbf{W})^{-1}[(\mathbf{I}-\rho\mathbf{W})']^{-1}\end{aligned} \tag{52}$$

したがって，誤差項に空間的自己相関がみられる場合の一般化最小2乗法は，式(45)，(46)，(47)に式(51)を代入したものになる．

このような一般化最小2乗法の利用は，Upton & Fingleton が指摘するように，計算が比較的容易であることや各種の変数値の回帰モデルに適用が可能であることなどから，さまざまな問題状況に対して行なうことができる（Upton & Fingleton, 1985, p.p. 277～283)．しかし，注意すべきは，この方法の利用に先立って，上述の行列$\mathbf{D}$が完全に知られていなければならないことである．このことは，前述のようにρの推定法が提唱されているとはいえ，Hepple (1976) が論ずるように，ほとんど不可能に近い．したがって，一般化最小2乗法の利用はこの意味ではかなり限定的であるといえる．このことを解決するためには，次に述べる最尤法を利用することができる．

4. 最　尤　法

1)　基 本 原 理

最尤法は，回帰パラメータの推定法としては最小2乗法とともにきわめて魅力のある方法であるといわれている[9]．それは，この方法が，データが与えられているとき，あるパラメータはそのデータをもっともらしくさせる値を取るべきである，という考えに基づいているからである．この考え方は，データにおけるY値をいくつかのパラメータに割り当てたとき，その割り当ての方式が何通りもあり，その方式を，もっともらしさを表現する尤度によって示すことで具体化される．この尤度全体，つまり結合尤度が最大になることは，パラメータがデータを最ももっともらしくさせる値を取ることに他ならないのである．したがって，パラメータの望ましい推定値は結合尤度を最大にさせるものであるといえる．最尤法は，最大の結合尤度をもたらすようなパラメータ推定値を求める方法である．

この方法の基本的前提は，一般的には，問題対象とされた事象に関する確率分布について想定されている前提であり，回帰モデルに関しては，従属変数，換言すれば誤差eの確率分布に対しての前提である．回帰モデルのeの確率分布は，前述のように，正規であると前提されている．

一般に，正規分布の確率密度関数は衆知のように次のように記されている．

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \tag{53}$$

ここでのμとσはXの平均と標準偏差である．したがって，サンプル地区iでの誤差e_iが正規分布

するとすれば，それの確率密度関数は次のようになる．

$$f(e_i)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-e_i^2}{2\sigma^2}\right) \tag{54}$$

ここでの σ は e の標準偏差である．n 個のサンプル地区全体を考え，e_i が独立かつ均等に分布すると仮定すれば，結合尤度を表わす関数 L は，次のような個々の密度関数の積になる．

$$L=\prod_{i=1}^{n}\frac{1}{\sigma\sqrt{2\pi}}\exp\left(\frac{-e_i^2}{2\sigma^2}\right) \tag{55}$$

または，$$L=\frac{1}{\sigma^n(2\pi)^{n/2}}\exp\left(\frac{-\mathbf{e}'\mathbf{e}}{2\sigma^2}\right) \tag{55'}$$

通常は，この関数が次のように対数変換された対数尤度関数が利用される．

$$\ln L=-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}(\mathbf{e}'\mathbf{e}) \tag{56}$$

$$\ln L=\text{constant}-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}(\mathbf{e}'\mathbf{e}) \tag{56'}$$

この式(56′)を回帰モデルの点から改書すると，次のようになる．

$$\ln L=\text{constant}-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}(\mathbf{Y}'\mathbf{Y}-2\boldsymbol{\beta}'\mathbf{X}'\mathbf{Y}+\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \tag{57}$$

回帰モデルの関する最尤法の問題は，$\mathbf{Y}$ と $\mathbf{X}$ を与えるとき，上式の $\ln L$ を最大にするような誤差 e の分散 σ^2 と回帰パラメータ $\boldsymbol{\beta}$ を推定することである．この方法で推定される最尤推定量は，いくつかの望ましい性質をもっている．それは，多くの場合，推定量が不偏であり，比較的小さい分散をもち，サンプル規模が大であるならば，それの分布が近似的に正規になるということである (Norden, 1972・1973).

2)　空間的自己相関に関連する最尤法

空間的自己相関を示す誤差項を伴う回帰モデルに関しては，上述の諸式は独立的な誤差を前提とした場合のものであるので，それらを修正する必要があることはいうまでもない．前節で述べたような $\mathbf{e}=(\mathbf{I}-\rho\mathbf{W})\mathbf{u}=\mathbf{A}\mathbf{u}$（これ以後 $(\mathbf{I}-\rho\mathbf{W})=\mathbf{A}$ とおく）なる自己回帰形式の相関があるとすれば，尤度関数は次のようになるであろう．

$$L=|\mathbf{A}|\frac{1}{\sigma^n(2\pi)^{n/2}}\exp\left[\frac{-(\mathbf{A}\mathbf{u})'\mathbf{A}\mathbf{u}}{2\sigma^2}\right] \tag{58}$$

この式が式 (55′) と大きく異なる点は，$\mathbf{A}$ の行列式 $|\mathbf{A}|$ が加えられていることである．これは，一般にヤコビ行列式と呼ばれる変換因であり，上式の $\mathbf{u}$ を $\mathbf{e}$ に変換させる因子である．式(58)を対数変換すると，次のようになる．

$$\ln L=\text{constant}-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}(\mathbf{u}'\mathbf{A}'\mathbf{A}\mathbf{u})+\ln|\mathbf{A}| \tag{59}$$

この式(59)を回帰モデルの点から改書すると，次のようになる．

$$\ln L=\text{constant}-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}(\mathbf{Y}'\mathbf{A}'\mathbf{A}\mathbf{Y}-2\boldsymbol{\beta}'\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{Y}$$
$$+\boldsymbol{\beta}'\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{X}\boldsymbol{\beta})+\ln|\mathbf{A}| \tag{60}$$

ここで注意すべきは，この式では推定されるべきものは σ^2 と $\boldsymbol{\beta}$ に加えて，$\mathbf{A}$ の中に含まれている ρ

の合計3つのパラメータであることである．

上式に関して，$\mathbf{Y}$と$\mathbf{X}$のデータ値を与えて，$\ln L$を最大にするようなσ^2と$\boldsymbol{\beta}$とρの推定値を求める問題は，直覚的に，求めるべき3つのパラメータに関して$\ln L$をそれぞれ偏微分して偏導関数を導き，それをゼロとおいた尤度方程式を連立方程式化すれば解きうるであろうと考えられる．

ρが既知の場合は（式(32)を思い出されたい），$\boldsymbol{\beta}$とσ^2の最尤推定量は，$\boldsymbol{\beta}$とσ^2に関する$\ln L$の偏導関数をゼロとおくことによって次のように得られる．すなわち，偏導関数は次式のとおりであり，

$$\frac{\partial(\ln L)}{\partial\boldsymbol{\beta}}=-\frac{1}{2\sigma^2}(-2\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{Y}+2\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{X}\boldsymbol{\beta}) \tag{61}$$

$$\frac{\partial(\ln L)}{\partial\sigma^2}=-\frac{n}{2\sigma^2}+\frac{1}{2\sigma^4}(\mathbf{e}'\mathbf{e}) \tag{62}$$

それゆえ，推定方程式は次のようになる．

$$\widehat{\boldsymbol{\beta}}=(\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{X})^{-1}\,\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{Y} \tag{63}$$

$$\widehat{\sigma^2}=\frac{1}{n}(\mathbf{Y}'\mathbf{A}'\mathbf{A}\mathbf{Y}-2\boldsymbol{\beta}'\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{Y}+\boldsymbol{\beta}'\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{X}\boldsymbol{\beta}) \tag{64}$$

ここで注目すべきは，前節において$\mathbf{A}'\mathbf{A}=\mathbf{D}^{-1}$とおいたが（式（51）を参照されたい），この$\mathbf{D}^{-1}$を式(63)に対して用いると，この式は前掲の式(45)と全く同等になることである．

ρが未知の場合は，このパラメータが式(60)にみられる$\mathbf{A}$の一部を構成しているので，ρに関する$\ln L$の尤度方程式を，$\boldsymbol{\beta}$とσ^2それぞれに関する尤度方程式を連立方程式化することによって，ρがσ^2と$\boldsymbol{\beta}$とともに同時に推定できると考えられる．これを達成しうる方法の1つが Doreian (1980) によって提示されている．式(64)を式(60)に代入すると，式(60)は次式のように簡単化される．

$$\ln L=\text{constant}-\frac{n}{2}\ln\widehat{\sigma^2}+\ln|\mathbf{A}| \tag{65}$$

そうすると，この式の$\ln L$のρに関する偏導関数から，$\widehat{\sigma^2}$と$\widehat{\boldsymbol{\beta}}$が既知の場合の$\widehat{\rho}$が得られる．この手続きは，次式を最小化することと同等である．

$$M=\ln\widehat{\sigma^2}-\frac{n}{2}\ln|\mathbf{A}| \tag{66}$$

ここにおいて，$\widehat{\sigma^2}=(\mathbf{Y}'\mathbf{A}'\mathbf{P}\mathbf{A}\mathbf{Y})$とする．$\mathbf{P}$は次式によって与えられるものである．

$$\mathbf{P}=\mathbf{I}-(\mathbf{A}\mathbf{X})\,[(\mathbf{A}\mathbf{X})'\,\mathbf{A}\mathbf{X}]^{-1}\,(\mathbf{A}\mathbf{X})' \tag{67}$$

したがって，式(66)を最小化することは次式を最小化することになる．

$$M_*=\ln(\mathbf{Y}'\mathbf{A}'\mathbf{P}\mathbf{A}\mathbf{Y})-\frac{n}{2}\ln|\mathbf{A}| \tag{68}$$

この式を用いれば，$\boldsymbol{\beta}$とσ^2とは無関係にρの推定量を導くことができるであろう．

以上の説明から，回帰分析において必要とされる$\boldsymbol{\beta}$とσ^2とρの3推定量を求める手順は，初めに式(68)に基づいて$\widehat{\rho}$を得，次いでこれを式(63)に投入することによって$\widehat{\boldsymbol{\beta}}$を求め，最後にその結果を式(64)に利用することによって$\widehat{\sigma^2}$を導く，というようになると理解されよう．しかし，3つのパラメータの推定量は式(60)から明らかなように相互に関連しており，それぞれが個別に導かれるものではない．厳密な推定量が必要とされる場合は，3つのパラメータが同時に推定できる枠組，つまり3パラメータそれぞれに関する$\ln L$の尤度方程式からなる連立方程式を解くという枠組が必要になるであろう．しかし，Hepple (1976) が指摘するように，3パラメータを同時に推定する方法は，2パラ

メータを推定する方法よりはるかに複雑となり，式 (63) や (64) のような推定式を導きえないのである．その場合は，3つの尤度方程式からなる連立方程式を解くことになるが，$n \times n$ の行列の逆行列を伴うここでの状況のような場合ではその解法計算は煩瑣になり，また長時間を要する．

3) 計 算 法

このようなことから，多数の未知数を推定するため考案されている非線形最適化法の利用が行なわれる．非線形最適化法にはさまざまな方法があり，最適解を直接解析的に求める方法（上述の2パラメータに関する最尤法はこれに当たる），モンテカルロ法を用いる確率的方法，近似解を逐次改良する逐次近似法に大別されるが（OR 事典編集委員会編，1975, p. 84), Hepple (1976) は，多数の未知数（ここでは3パラメータ）を，短時間にしかもわずかな誤差で推定する方法として，Powell (1964) によって考案された共役方向 (conjugate direction) アルゴリズムの利用を薦めている[10]．

Powell の方法は，最小化または最大化さるべき目的関数（尤度関数）の尤度方程式を使用しないことに特徴があり，2次関数に関するベクトルの共役性を巧みに利用している．一般に，2次関数の2階偏導関数行列 $\mathbf{G}$ が対称であり，k 個のベクトル $\mathbf{d}^{(1)}, \mathbf{d}^{(2)}, \cdots, \mathbf{d}^{(k)}$ がある場合，次の条件を満たすとき，

$$\mathbf{d}^{(i)\prime}\mathbf{G}\mathbf{d}^{(j)}=0 \qquad (i \neq j) \tag{69}$$

k 個のベクトルは $\mathbf{G}$ に関して互いに共役であるといわれる[11]．これに関して1つの基本的定理がある．それは，「$\mathbf{G}$ が対称行列であるならば，$\mathbf{G}$ に関して互いに共役な n 個のゼロでないベクトルが存在し，さらに $\mathbf{G}$ が正値定符号であるならば，$\mathbf{G}$ に関して互いに共役なゼロでないベクトルは1次独立である」である．そして，2次関数の最小値を見出すための探索において「n 個の共役な方向の各々に沿って探索を順次行なうならば，2次関数の最小点が得られる」という定理がある．Powell の方法を初めとする多くの共役方向法は，これらの2つの定理を基礎にしており，それら以外の付加的な定理や計算効率に関する概念などを導入することによって，さまざまな工夫が施されている[12]．

Powell の方法のアルゴリズムはかなり複雑であり，推定すべき未知数が多い場合は計算反復回数が増大すること，また，その過程で新しい共役方向ベクトルが仲々うまく採用されないことなどの難点があるが，多くの共役方向法のうちで最も有効なものの1つであると評価されている（今野・山下, 1978). Powell 法のアルゴリズムは次のとおりである[13]．

ステップ0：1次独立的な共役方向の初期ベクトル $\mathbf{d}_0^{(0)}$，$\mathbf{d}_1^{(0)}, \cdots, \mathbf{d}_{n-1}^{(0)}$（$n$ は未知数の個数）と推定すべき未知数の初期ベクトル $\mathbf{x}_B^{(0)}$ を用意し，$k=0$ とする．

ステップ1：$\mathbf{x}_0^{(k)}=\mathbf{x}_B^{(k)}$ とおき，次式で表わされる $\mathbf{x}_{i+1}^{(k)}$ を $i=0, 1, \cdots, n-1$（i は探索方向）について求める．

$$\mathbf{x}_{i+1}^{(k)}=\mathbf{x}_i^{(k)}+\alpha_i^{(k)}\,\mathbf{d}_i^{(k)} \in \alpha_i^{(k)}=\min_{\alpha} f(\mathbf{x}_i^{(k)}+\alpha\,\mathbf{d}_i^{(k)}) \tag{70}$$

ステップ2：$\mathbf{d}_n^{(k)}=\mathbf{x}_n^{(k)}-\mathbf{x}_0^{(k)}$ とおいて，次式から $\mathbf{x}_{n+1}^{(k)}$ を求める．

$$\mathbf{x}_{n+1}^{(k)}=\mathbf{x}_{n}^{(k)}+\alpha_{n}^{(k)}\ \mathbf{d}_{n}^{(k)} \in \alpha_{n}^{(k)}=\min_{\alpha} f(\mathbf{x}_{n}^{(k)}+\alpha\ \mathbf{d}_{n}^{(k)}) \tag{71}$$

ステップ３：$\mathbf{x}_{B}^{(k+1)}=\mathbf{x}_{n+1}^{(k)}$ とおく．

ステップ４：次式が成立するならば，作業終了，$\mathbf{x}_{B}^{(k)}$ が最適解．成立しないときは次ステップを行なう．

$$\| \mathbf{x}_{B}^{(k+1)}-\mathbf{x}_{B}^{(k)} \| \approx 0 \tag{72}$$

ステップ５：次式によって p を見定める．

$$f(\mathbf{x}_{p}^{(k)})-f(\mathbf{x}_{p+1}^{(k)})=\max_{0\leq i\leq n-1}[f(\mathbf{x}_{i}^{(k)})-f(\mathbf{x}_{i+1}^{(k)})] \tag{73}$$

そして，次式が満たされるか否かを判別する．

$$|\alpha_{n}^{(k)}|<\left[\frac{f(\mathbf{x}_{n}^{(k)})-f(\mathbf{x}_{n+1}^{(k)})}{f(\mathbf{x}_{p}^{(k)})-f(\mathbf{x}_{p+1}^{(k)})}\right] \tag{74}$$

満たされる場合は，$\mathbf{d}_{i}^{(k+1)}=\mathbf{d}_{i}^{(k)}$ とし，満たされない場合は，

$$\mathbf{d}_{i}^{(k+1)}=\begin{cases}\mathbf{d}_{i}^{(k)} & (i=0,\cdots,p-1)\\ \mathbf{d}_{i+1}^{(k)} & (i=p,\cdots,n-1)\end{cases}$$

として，$k=k+1$ とおいてステップ１へゆく[14]．

なお，最小化・最大化問題のコンピュータプログラムが GLIM や GENSTAT のパッケージにあるといわれているが，どのようなアルゴリズムに拠るのかは未詳である[15]．

4）推　　測

上述のような手続きによって回帰パラメータの最尤推定量が得られると，この推定量からなる回帰モデルの適合性の検定およびパラメータ推定量の有意性の検定が必須の作業となる．この作業に際しては，通常の最小２乗法推定量の場合は，前者に関しては重決定係数および後者に関しては t 統計量がそれぞれ利用されるが，最尤推定量の場合は，それらを直截的には適用しえない．とりわけ，t 統計量についてはそうである．

回帰モデルの適合性の検討については，「回帰モデルの適合性は，それがどのようなものであっても，回帰分散の全分散に対する比によって表現される」という考え方が，ここの場合においても基本原理であり，通常の重決定係数ときわめて類似した次式を利用する．

$$r^{*2}=\frac{\sum_{i=1}^{n}(y_i-\widehat{y})^2/n-\widehat{\sigma^2}}{\sum_{i=1}^{n}(y_i-\widehat{y})^2/n} \tag{75}$$

この式の $\widehat{\sigma^2}$ が最小２乗推定量であるならば，この式は通常の重決定係数 (r^2) と同等になる．

回帰パラメータの最尤推定量の有意性検定については．最尤推定量が漸近（大標本）的に正規性と独立性を有するということを前提として行なわれ，（各パラメータの推定値／それの標準誤差）と正規確率限界値とが比較される．この際に問題となることは，各パラメータの標準誤差をどのように求めるかである．この標準誤差は，Kendall & Stuart が提示したパラメータに関する分散共分散行列

$\mathbf{V}$の中に含まれているので (Kendall & Stuart, 1961, p. 51), この分散共分散行列を求めればよいことになる．この行列は次のように表わされ，行列の各要素は対数尤度関数の各パラメータに関する2階偏導関数の期待値である．

$$\mathbf{V}=\left[\frac{-\mathrm{E}\,(\partial^2 \ln L)}{\partial\theta_r\partial\theta_s}\right]^{-1} \tag{76}$$

ここでの θ_r と θ_s はそれぞれ r 番目と s 番目のパラメータである．回帰モデルにおけるパラメータは $\boldsymbol{\beta}$ と σ^2 と ρ の3つであるが，そのうちの $\boldsymbol{\beta}$ と σ^2 に関する1階偏導関数はそれぞれ前掲の式(61)と(62)のとおりであるが，ρ に関するそれは次式で与えられる．

$$\frac{\partial(\ln L)}{\partial\rho}=-\sum_{i=1}^{n}\frac{\lambda_i}{1-\rho\lambda_i}+\frac{1}{\sigma^2}\mathbf{V}'\mathbf{W}\mathbf{u} \tag{77}$$

ここでの λ_i は行列 $\mathbf{W}$ の i 番目の固有値, $\mathbf{V}$は式(52)で定義される誤差項の分散共分散行列 $(=\sigma^2[\mathbf{A}'\mathbf{A}]^{-1})$ である．上式および式(61)と(62)に基づいて導出される2階偏導関数の期待値は次のとおりである．

$$\mathrm{E}\left[\frac{\partial^2(\ln L)}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}}\right]=-\frac{1}{\sigma^2}\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{X} \tag{78}$$

$$\mathrm{E}\left[\frac{\partial^2(\ln L)}{\partial\sigma^2\partial\sigma^2}\right]=-\frac{1}{2\sigma^4}\,n \tag{79}$$

$$\mathrm{E}\left[\frac{\partial^2(\ln L)}{\partial\rho\partial\rho}\right]=\alpha-\mathrm{tr}(\mathbf{B}'\mathbf{B}) \tag{80}$$

$$\mathrm{E}\left[\frac{\partial^2(\ln L)}{\partial\sigma^2\partial\rho}\right]=-\frac{1}{\sigma^2}\mathrm{tr}(\mathbf{B}) \tag{81}$$

ここでの$\alpha=-\sum_{i=1}^{n}\lambda_i^2/(1-\rho\lambda_i)^2$, $\mathbf{B}=\mathbf{W}\mathbf{A}^{-1}$ である．上式で示される4つの2階偏導関数以外の2つのもの，つまり $\mathrm{E}\,[\partial^2(\ln L)/\partial\boldsymbol{\beta}\partial\rho]$ と $\mathrm{E}\,[\partial^2(\ln L)/\partial\boldsymbol{\beta}\partial\sigma^2]$ はゼロである．かくして，3つの回帰パラメータ $\boldsymbol{\beta}$, σ^2, ρ の最尤推定量に関する分散共分散行列は次のように示される．

$$\mathbf{V}(\widehat{\sigma^2},\hat{\rho},\hat{\boldsymbol{\beta}})=\sigma^4\begin{bmatrix} n/2 & \sigma^2\mathrm{tr}(\mathbf{B}) & 0 \\ \sigma^2\mathrm{tr}(\mathbf{B}) & \sigma^4\,\mathrm{tr}\,(\mathbf{B}'\mathbf{B})-\sigma^4\alpha & 0 \\ 0 & 0 & \sigma^2\mathbf{X}'\mathbf{A}'\mathbf{A}\mathbf{X}\end{bmatrix}^{-1} \tag{82}$$

この行列の主対角要素の平方根が，各パラメータの推定量の漸近標準誤差である．したがって，上式(82)にデータ数値を投入すれば，検定のための標準誤差が得られる．

上述の有意性検定において利用される$\mathbf{W}$の固有値 λ について注目すべき事柄が2つある．その1つは，$\mathbf{A}=(\mathbf{I}-\rho\mathbf{W})$ とおいていることから，次式が成立することである．

$$|\mathbf{A}|=\prod_{i=1}^{n}(1-\rho\lambda_i) \tag{83}$$

$$\ln|\mathbf{A}|=\sum_{i=1}^{n}\ln(1-\rho\lambda_i) \tag{84}$$

つまり，n 個の λ_i の積は $|\mathbf{A}|$ を与えるということである．前掲の式(60)の $\ln L$ を最大にするような3つのパラメータのうちの ρ の推定量は，式(84)を式(60)に代入して求められ，また，式(77)の導出もこれに拠っているのである．いま1つは，サンプル地区数 n が大な場合は固有値を求めることが煩瑣になるが, Ord (1975) によって，ある状況下でこれを簡単に入手できる方法が提示されているこ

とである．さまざまな状況下での方法が示されているが，そのうちで最も利用度の高いものは次のとおりである．p 行 q 列の四角形メッシュ ($n=p\times q$) を想定し，そのメッシュにおける行 r 列 s に当たるメッシュ方格での変数 y_{rs} が，次に示されるような1階の空間的自己相関を呈するとする（ただし，メッシュの境界に接する方格での変数値は調整されているとする）．

$$y_{rs}=\frac{1}{4}\rho(y_{r-1,s}+y_{r+1,s}+y_{r,s-1}+y_{r,s+1})+e_{rs} \tag{85}$$

このような状況下での行列 **W** は $pq\times qp(=n\times n)$ の対称行列になるが，メッシュでの行 r 列 s に当たる方格に関する**W**の固有値 λ_{rs} は次式で与えられる．

$$\lambda_{rs}=\frac{1}{2}\{\cos[r\pi/(p+1)]+\cos[s\pi/(q+1)]\} \tag{86}$$

回帰パラメータの最尤推定量の有意性検定においては，上述のような標準誤差を利用する方法の他に，尤度比を利用する方法もある．この方法は，概述すれば，$-2\ln(L_R/L_U)=2\ln(L_U/L_R)$ がカイ2乗分布に従うことに基づいている．L_U は非制約モデルの尤度関数であり，L_R は制約モデルのそれである．ここでの非制約モデルは，すべての最尤推定量を含む回帰モデルに関する尤度関数である（式 (60) を参照されたい）．後者の制約モデルについては，m 個の回帰パラメータの中の k 番目のパラメータ β_k に関して帰無仮説 $H_0:\beta_k=0$（対立仮説 $H_1:\beta_k\neq 0$）を検定しようとする場合は，$\widehat{\beta}_k=0$ としたときの回帰モデルに関する尤度関数がこれに該当する．

上述の回帰パラメータの最尤推定量の有意性検定の2方法の他に，非制約モデルの対数尤度関数の $\boldsymbol{\beta}$ に関する2階偏導関数のみを問題にする Wald 統計量による検定法や[16]，制約モデルの対数尤度関数の $\boldsymbol{\beta}$ に関する1階と2階の両方の偏導関数を問題にする Lagrange 乗数による検定法などもある (Pickles, 1986). Wald 統計量および Lagrange 乗数ともにカイ2乗分布に従う．

ここで注目すべきは，同一の推測問題に各種の検定法を適用すると，その結果は検定法によって少しずつ異なることである．どの検定法が最適であるかは未解決の問題である．

注

1）このような影響を最初に指摘したのは Student (1914) であり，彼は，今日でいう傾向面分析によってデータ数値から自己相関成分を除去し，その結果値による平均の推定法を提案している．

2）Geary (1954) にこれの原図が記載されている．

3）$(\rho_1, \rho_2)=(0.0, 0.0)$ の場合は，2変数ともに自己相関をもたないことを意味し，2変数それぞれが独立的に分布する状況である．

4）例えば，Durbin & Watson (1950・1951・1971), Christ (1966), Johnston (1972) などがあげられる．

5）Johnston (1972) の p.p. 18〜19 を参照されたい．

6）$\widehat{\boldsymbol{\beta}}=(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}\mathbf{Y}=\boldsymbol{\beta}+(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$ なることを利用する．

7）正規分布の仮定のもとで，I_k の平均と分散は次のように導かれている．

$$E(I_k)=\mathrm{tr}(\mathbf{MW})/(n-m)$$

$$\mathrm{var}(I_k)=\frac{\mathrm{tr}(\mathbf{MWMW}')+\mathrm{tr}(\mathbf{MWMW})+[\mathrm{tr}(\mathbf{MW})]^2}{(n-m)(n-m+2)}-E(I_k)^2$$

ここでの $\mathbf{M}=\mathbf{I}-\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}$. Cliff & Ord (1972) もまた次の統計量を導いている．

$$I=(n/W)\sum_{j=1}^{n}\sum_{i=1}^{n}w_{ij}\,\widehat{e}_i\,\widehat{e}_j\Big/\sum_{i=1}^{n}\widehat{e}_i^{\,2}$$

ここでの $W=\sum_{j=1}^{n}\sum_{i=1}^{n}w_{ij}$ である．

8）詳事は奥野（1981）を参照されたい.

9）回帰パラメータの推定問題を含む各種の地理学的問題に対する最尤法の利用が，Pickles（1986）によって平易に解説されている．また，最尤法の理論については福原（1981）の p.p. 37～59 に詳述されている.

10) Powell (1964) によって考案されたアルゴリズムに対してさまざまな解釈が行なわれており，擬似ニュートン法の一種（OR 事典編集委員会編，1975），共役傾斜法の一種 (Kowalik & Osborne, 1968)，直接探索法の１つである共役方向法 (Dixon, 1972) などと解されている．Hepple は共役傾斜法としているが，後述するアルゴリズムからみれば，共役方向法と解すべきであろう.

11) G＝I のときは $\mathbf{d}^{(i)\prime}\mathbf{d}^{(j)}=0$ になり，$\mathbf{d}^{(i)\prime}$ と $\mathbf{d}^{(j)}$ はGに関して互いに直交関係にあることになる．したがって，共役は直交の概念を一般化したものといえる.

12) このことについては，Dixon (1972) の p.p. 49～87 およびその訳礬松原訳（1974）の p.p. 53～97と今野・山下（1978）の p.p. 169～195 を参照されたい.

13) ここで述べるアルゴリズムは，原著のものを平易な形で整理された今野・山下（1978）の p.p. 178～182に準拠している.

14) このアルゴリズムの理解のため，１つの計算例を示す．最小化すべき目的関数は次のとおりであり，推定すべき未知数は x_1, x_2, x_3 の３つである.

$$f(x_1, x_2, x_3)=(x_1-x_2+x_3)^2+(-x_1+x_2+x_3)^2+(x_1+x_2-x_3)^2$$

この関数の２階偏導関数行例Gは次のように得られる.

$$\mathrm{G}=\begin{pmatrix}6 & -2 & -2\\ -2 & 6 & -2\\ -2 & -2 & 6\end{pmatrix}$$

ステップ０：未知数初期ベクトル $\mathbf{X}_B^{(0)}$ および共役方向初期ベクトル $\mathbf{d}_0^{(0)}$，$\mathbf{d}_1^{(0)}$，$\mathbf{d}_2^{(0)}$ として次のようにおくことにする (Powell は共役方向初期ベクトルとして単位ベクトルを薦めている).

$\mathbf{x}_B^{(0)}=(1/2, 1, 1/2)'$（この場合は $f(\mathbf{x}_0^{(0)})=2$ となる），

$\mathbf{d}_0^{(0)}=(1, 0, 0)'$, $\mathbf{d}_1^{(0)}=(0, 1, 0)'$, $\mathbf{d}_2^{(0)}=(0, 0, 1)'$

ステップ１：

① $\mathbf{d}_0^{(0)}$ 方向の直線探索を行なう．$\mathbf{x}_0^{(0)}=\mathbf{x}_B^{(0)}$ とおいて，初めに式(70)にみられるような，$f(\mathbf{x}_0^{(0)}+\alpha\mathbf{d}_0^{(0)})$ を最小化する $\alpha_0^{(0)}$ を，次のような作業で見出す.

$$\begin{aligned}f(\mathbf{x}_0^{(0)}+\alpha\mathbf{d}_0^{(0)})&=f[(1/2, 1, 1/2)'+\alpha(1, 0, 0)']\\ &=f(1/2+\alpha, 1, 1/2)'=(1/2+\alpha-1+1/2)^2\\ &\quad+(-1/2-\alpha+1+1/2)^2+(1/2+\alpha+1-1/2)^2\\ &=2+3\alpha^2\end{aligned}$$

$$\partial f/\partial\alpha=6\alpha=0 \qquad \therefore \quad \alpha=0$$

次に，ここで見出された α に基づいて式(70)から $\mathbf{x}_1^{(0)}$ を次のように求める.

$$\mathbf{x}_1^{(0)}=\mathbf{x}_0^{(0)}+\alpha_0^{(0)}\mathbf{d}_0^{(0)}=(1/2, 1, 1/2)'+0\times(1, 0, 0)'=(1/2, 1, 1/2)'$$

このときの $f(\mathbf{x}_1^{(0)})=2$ であり，$f(\mathbf{x}_0^{(0)})$ と同等の関数値を取る.

② $\mathbf{d}_1^{(0)}$ 方向の直線探索を行なう．①の場合と同様な作業で $f(\mathbf{x}_1^{(0)}+\alpha\mathbf{d}_1^{(0)})$ を最小化する $\alpha_1^{(0)}$ を見出す.

$$f(\mathbf{x}_1^{(0)}+\alpha\mathbf{d}_1^{(0)})=2+4\alpha+3\alpha^2$$

$$\partial f/\partial\alpha=4+6\alpha=0 \qquad \therefore \quad \alpha=-2/3$$

したがって，

$$\mathbf{x}_2^{(0)}=\mathbf{x}_1^{(0)}+\alpha_1^{(0)}\mathbf{d}_1^{(0)}=(1/2, 1, 1/2)'-2/3\,(0, 1, 0)'=(1/2, 1/3, 1/2)'$$

このときの $f(\mathbf{x}_2^{(0)})=2/3$ であり，上述の $f(\mathbf{x}_1^{(0)})$ より 4/3 だけ減少している.

③ $\mathbf{d}_2^{(0)}$ 方向の直線探索を行なう．①と②の場合と同様にして $f(\mathbf{x}_2^{(0)}+\alpha\mathbf{d}_2^{(0)})$ を最小化する $\alpha_2^{(0)}$ を見出す.

$$f(\mathbf{x}_2^{(0)}+\alpha\mathbf{d}_2^{(0)})=6/9+(4/3)\alpha+3\alpha^2$$

$$\partial f/\partial\alpha=4/3+6\alpha=0 \qquad \therefore \quad \alpha=-2/9$$

したがって，

$$\mathbf{x}_3^{(0)}=\mathbf{x}_2^{(0)}+\alpha_2^{(0)}\mathbf{d}_2^{(0)}=(1/2, 1/3, 1/2)'-2/9\,(0, 0, 1)'=(1/2, 1/3, 5/18)'$$

このときの $f(\mathbf{x}_3^{(0)})=14/27$ であり，$f(\mathbf{x}_2^{(0)})$ より4/27だけ減少している.

以上によって，初めに設定された未知数ベクトルに関する3方向それぞれについての探索が終了し，$f(\mathbf{x}_0^{(0)})$ より小の関数値をもたらす未知数ベクトル $\mathbf{x}_3^{(0)}$ があることが判明した．換言すれば，３次元空間での初期座標点 $\mathbf{x}_0^{(0)}$ が３座標軸に沿って新しい座標点 $\mathbf{x}_3^{(0)}$ に移動したのである．そこで，この座標点が最適解であるか否かを検討するとともに，次段階の探索の準備を行なう.

ステップ２：$\mathbf{d}_3^{(0)}=\mathbf{x}_3^{(0)}-\mathbf{x}_0^{(0)}=(1/2, 1/3, 5/18)'-(1/2, 1, 1/2)'=(0, -2/3, -2/9)'$ とおき，$f(\mathbf{x}_3^{(0)}+$

$\alpha \mathbf{d}_3^{(0)}$) を最小化する $\alpha_3^{(0)}$ を見出す.

$f(\mathbf{x}_3^{(0)}+\alpha\mathbf{d}_3^{(0)})=14/27-(8/27)\alpha+(32/27)\alpha^2$

$\partial f/\partial\alpha=-8/27+(64/27)\alpha=0 \qquad \therefore \quad \alpha=1/8$

したがって,

$\mathbf{x}_4^{(0)}=\mathbf{x}_3^{(0)}+\alpha_3^{(0)}\mathbf{d}_3^{(0)}=(1/2,\ 1/3,\ 5/18)'$
$\quad +1/8\,(0,\ -2/3,\ -2/9)'=(1/2,\ 1/4,\ 1/4)'$

このときの $f(\mathbf{x}_4^{(0)})=1/2$ であり, $f(\mathbf{x}_3^{(0)})$ に比べてわずか 1/54 だけ減少しているにすぎない.

ステップ 3 : $\mathbf{x}_B^{(1)}=\mathbf{x}_4^{(0)}=(1/2,\ 1/4,\ 1/4)'$ とおく.

ステップ 4 : 式(72)が成立するか否かを検討する.

$\|\mathbf{x}_B^{(1)}-\mathbf{x}_0^{(0)}\|=\|(1/2,\ 1/4,\ 1/4)'-(1/2,\ 1,\ 1/2)'\|$
$\quad =(0,\ 3/4,\ 1/4)'\fallingdotseq 0$

したがって, 最適解が得られていないことが判明する. そこでステップ 5 へ進む.

ステップ 5 : 式(73)の右辺を調べることによって p を見定める.

$f(\mathbf{x}_0^{(0)})-f(\mathbf{x}_1^{(0)})=2-2=0$

$f(\mathbf{x}_1^{(0)})-f(\mathbf{x}_2^{(0)})=2-2/3=4/3$

$f(\mathbf{x}_2^{(0)})-f(\mathbf{x}_3^{(0)})=2/3-14/27=4/27$

$f(\mathbf{x}_3^{(0)})-f(\mathbf{x}_4^{(0)})=14/27-1/2=1/54$

であるので, 各対(つい)の関数値の最大差は, $f(\mathbf{x}_1^{(0)})-f(\mathbf{x}_2^{(0)})=4/3$ のものである. それゆえ, $p=1$ と見定められる. そこで, 次に式(74)の不等式を調べる. その結果は,

$$\left[\frac{f(\mathbf{x}_3^{(0)})-f(\mathbf{x}_4^{(0)})}{f(\mathbf{x}_1^{(0)})-f(\mathbf{x}_2^{(0)})}\right]^{1/2}=\left(\frac{1/54}{4/3}\right)^{1/2}=\frac{1}{\sqrt{72}}$$

と得られ, これは $\alpha_3^{(0)}=1/8$ より小である. それゆえ, 式(74)は満たされない. そこで, $\mathbf{d}_0^{(1)}=\mathbf{d}_0^{(0)}=(1,0,0)'$, $\mathbf{d}_1^{(1)}=\mathbf{d}_2^{(0)}=(0,0,1)'$, $\mathbf{d}_2^{(1)}=\mathbf{d}_3^{(0)}=(0,\ -2/3,\ -2/9)'$ として, ステップ 1 へゆく ($\mathbf{x}_4^{(0)}=\mathbf{x}_B^{(1)}=(1/2,\ 1/4,\ 1/4)'$. つまり $\mathbf{x}_4^{(0)}$が 2 段階目の作業の初期座標点になる). このことは, 2 段階目 ($k=1$) の探索を行なうに際して共役方向ベクトルが改められたことを意味する.

2 段階目の探索では, $\mathbf{x}_B^{(1)}=(1/2,\ 1/4,\ 1/4)'$ を出発点として, 上述の 3 つの共役方向ベクトルの直線探索を行なう. その結果は次のとおりである.

$\mathbf{x}_0^{(1)}=(1/2,\ 1/4,\ 1/4)'$, $f(\mathbf{x}_0^{(1)})=1/2$,

$\alpha_0^{(1)}=-1/3$

$\mathbf{d}_0^{(1)}$: $\mathbf{x}_1^{(1)}=(1/6,\ 1/4,\ 1/4)'$, $f(\mathbf{x}_1^{(1)})=1/6$,

$\alpha_1^{(1)}=-1/9$

$\mathbf{d}_1^{(1)}$: $\mathbf{x}_2^{(1)}=(1/6,\ 1/4,\ 5/36)'$, $f(\mathbf{x}_2^{(1)})=7/54$,

$\alpha_2^{(1)}=1/4$

$\mathbf{d}_2^{(1)}$: $\mathbf{x}_3^{(1)}=(1/6,\ 1/12,\ 1/12)'$, $f(\mathbf{x}_3^{(1)})=1/18$,

$\alpha_3^{(1)}=1/2$

$\mathbf{d}_3^{(1)}=(-1/3,\ -1/6,\ -1/6)'$, $\mathbf{x}_4^{(1)}=(0,\ 0,\ 0)'$,

$f(\mathbf{x}_4^{(1)})=0$

この場合では $p=0$ となり, $\{[f(\mathbf{x}_3^{(1)})-f(\mathbf{x}_4^{(1)})]/[f(\mathbf{x}_0^{(1)})-f(\mathbf{x}_1^{(1)})]\}^{1/2}=1/\sqrt{6}$ と求められる. これは $\alpha_3^{(1)}=1/2$ より小である. それゆえ, 3 段階目 ($k=2$) の探索に際しての共役方向ベクトルは, $\mathbf{d}_0^{(2)}=\mathbf{d}_1^{(1)}=(0,\ 0,\ 1)'$, $\mathbf{d}_1^{(2)}=\mathbf{d}_2^{(1)}=(0,\ -2/3,\ -2/9)'$, $\mathbf{d}_2^{(2)}=\mathbf{d}_3^{(1)}=(-1/3,\ -1/6,\ -1/6)'$ と改められる.

3 段階目の探索における出発点は $\mathbf{x}_B^{(2)}=(0,\ 0,\ 0)'$ であり, 改められた上述の 3 つの共役方向ベクトルの直線探索を行なうと, $\mathbf{x}_4^{(2)}=(0,\ 0,\ 0)'$ となり, ステップ 3・4 によれば, $\mathbf{x}_B^{(3)}=\mathbf{x}_4^{(2)}$ とおかれ, $\mathbf{x}_B^{(3)}$ は $\mathbf{x}_B^{(2)}(=\mathbf{x}_4^{(1)})$ と全く等しく, $(0,0,0)'$ である. したがって, 作業はこの段階で終了し, 得られた最適解は, $(x_1,\ x_2,\ x_3)=(0,0,0)$ である. この場合での関数値はゼロになる.

15) 1984年現在のイギリスにおいて入手しうるマイクロコンピュータ用のソフトウェアーパッケージについて, O'Brien (1986) が紹介している.

16) 回帰パラメータの最尤推定量の有意性検定に際しては, Wald 統計量の利用が比較的簡単であるといわれている (Pickles, 1986). それは, $\mathrm{H}_0 : \beta_k=0$ の場合, この統計量は$W=\hat{\beta}_k^2/\mathrm{var}\,(\hat{\beta}_k)$ に縮小され, $\mathrm{var}\,(\hat{\beta}_k)$ は式(82)に定義された行列にみられる $\hat{\beta}$ に関する主対角要素によって示されるからである.

参 考 文 献

ＯＲ事典編集委員会編（1975）:「ＯＲ事典」日科技連，569 p.

奥野隆史（1981）：空間的自己相関論（Ｉ）―測度と検定について―．人文地理学研究，5，165-183.

今野　浩・山下　浩（1978）：「非線形計画法」日科技連，354 p.

田中和子（1982）：空間的自己相関研究の展望とパターン検定の改良．地理学評論，55，313-333.

福原文雄（1981）：「計量経済分析の方法」マグロウヒル好学社，369 p.

Aitkin, A. C. (1934): On least-squares and linear combinations of observations. *Proceedings of the Royal Society of Edinburgh, Ser. A*, **55**, 42-47.

Brandsma, A. S. and Ketellapper, R. H. (1979): Further evidence on alternative procedures for testing of spatial autocorrelation among regression disturbances. Barteles, C. P. A. and Ketellapper, R. H. eds.: *Exploratory and explanatory statistical analysis of spatial data.* Martinus Nijhoff: Boston, 113-136.

Christ, C. (1966): *Econometric models and methods.* John Wiley & Sons: New York, 705 p.

Cliff, A. D. and Ord, J. K. (1972): Testing for spatial autocorrelation among regression residuals. *Geographical Analysis*, **4**, 267-284.

Cliff, A. D. and Ord, J. K. (1975a): The comparison of means when samples consist of spatially autocorrelated observations. *Environment and Planning, Ser. A*, **7**, 724-734.

Cliff, A. D. and Ord, J. K. (1975b): The choice of a test for spatial autocorrelation. Davis, J. C. and McCullagh, M. J. eds.: *Display and analysis of spatial data.* John Wiley & Sons: London, 54-77.

Cliff, A. D. and Ord, J. K. (1981): *Spatial processes: models & applications.* Pion: London, 266 p.

Dixon, L. C. W. (1972): *Nonlinear optimisation.* English Universities Press: London, 214 p. [松原正一訳（1974）:「非線形最適化計算法」培風館，243 p.]

Doreian, P. (1980): Linear models with spatially distributed data. *Sociological Methods and Research*, **9**, 29-60.

Durbin, J. and Watson, G. S. (1950・1951・1971): Testing for serial correlation in least-squares regression (I)・(II)・(III). *Biometrika*, **37**, 409-428, **38**, 159-178, **58**, 1-19.

Geary, R. C. (1954): The contiguity ratio and statistical mapping. *The Incorporated Statistician*, **5**, 115-145 [Berry, B. J. L. and Marble, D. F. eds. (1968): *Spatial analysis: a reader in statistical geography.* Prentice-Hall: Englewood Cliffs, 461-478].

Haggett, P., Cliff, A. D. and Frey, A. (1977): *Locational analysis in human geography.* 2nd ed. Edward Arnold: London, 605 p.

Haining, R. P. (1977): Model specification in stationary random fields. *Geographical Analysis*, **9**, 107-129.

Haining, R. P. (1978): The moving average model for spatial interaction. *Transactions of the Institute of British Geographers, New Series*, **3**, 202-225.

Hepple, L. W. (1976): A maximum likelihood model for econometric estimation with spatial series. Masser, I. ed.: *Theory and practice in regional science (London papars in regional science 6).* Pion: London, 90-104.

Johnston, J. (1972): *Econometric methods.* 2nd. ed. McGraw Hill: London, 437 p. [竹内 啓・関谷 章・栗山規矩・美添泰人・舟岡史雄訳（1975・1976）：「計量経済学の方法 全訂版」東洋経済新報社，523 p.]

Kendall, M. G. and Stuart, A. (1961): *The advanced theory of statistics. Vol. 2 Inference and relationship.* 3rd ed. Charles Griffin: London, 676 p.

Kowalik, J. and Osborne, M. R. (1968): *Methods for unconstrained optimization problems.* American Elsevier Pub.: New York, 148 p. [山本善之・小山健夫訳（1970）:「非線形最適化問題：制約条件のない最適化の手法」培風館，165 p.].

Martin, R. L. (1974): On spatial dependence, bias and the use of first spatial differences in regression analysis. *Area*, **6**, 185-194.

Norden. R. H. (1972・1973): A survey of maximum likelihood estimation. *International Statistical Review*, **40**, 329-354, **41**, 39-58.

O'Brien, L. G. (1986): Statistical software for microcomputers. *Area*, **18**, 39-42.

Ord, J. K. (1975): Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, **70**, 120-126.

Pickles, A. (1986): *An introduction to likelihood analysis (CATMOG 42).* Geo Books: Norwich, 48 p.

Powell, M. J. D. (1964): An efficient method for finding the minimum of function of several variables without calculating derivatives. *The Computer Journal*, 7, 155-162.

Student (1914): The elimination of spurious correlation due to position in time or space. *Biometrika*, **10**, 179-180.

Upton, G. J. G. and Fingleton, B. (1985): *Spatial data analysis by example. Vol. 1 Point pattern and quantitative data.* John Wiley & Sons: Chichester, 410 p.

Yule, G. U. (1926): Why do we sometimes get nonsense-correlations between time series? A study in sampling and the nature of time series. *Journal of the Royal Statistical Society*, **89**, 1-69.

Discussion on Spatial Autocorrelation

——Mainly on Its Noise Effect and Removal Methods——

Takashi OKUNO

Spatial autocorrelation is one of the most important and fundamental themes in quantitative geographical research. In fact, when A. D. Cliff and J. K. Ord reported some characteristics of modelling and analyzing in spatial context to statisticians in the joint meeting of the Royal Statistical Society and the Institute of British Geographers of 1975, they proposed that to develop this correlation was the one of four main tasks for the application of statistical concepts and methods to human geographic phenomena. They claimed that some aspects of spatial autocorrelation are obviously associated with an essential subject of geography and a much important technical problem of quantitative geography. In the previous paper (Okuno, 1981), the author reviewed the advantages and disadvatages of five measures and their statistical moments, beginning with P. A. P. Moran's proposal of 1948. This continued paper, based on more recent contributions of spatial autocorrelation, aims firstly to investigate various biases which spatial correlation in cross-sectional data takes carry into the statistical test and inferential result and secondly to review some methods to remove the biases. With the first point, two kinds of biases, which are regarded "noise effect" by traditional statisticians, are considered. These are for the *t*-test statistics used in testing the significant difference between two sample means, as shown in Cliff and Ord (1975a), and also for the ordinary least-squares estimator of linear regression parameter with spatially autocorrelated error term. With the second point, two ways of overcoming these problems are discussed from the sense of practically treating cross-sectional data. One is the spatial differencing method as a filter of the data to remove the spatial correlation in straightforward manner. Another is the modified *t*-statistics by Cliff and Ord (1975a) and the generalized least-squares and the maximum likelihood method, which produce the modified regression models allowing for the correlation. In the context of calculating maximum of the logarithmic likelihood function, the Powell's method (Powell, 1964), being one of the conjugate direction methods, particularly is recommended as the most effective algorithm.