

# 一般化 Whittle 法による不等間隔時空間データの 分析

松田 安昌<sup>†</sup>

(受付 2011 年 6 月 29 日；改訂 11 月 18 日；採択 12 月 22 日)

## 要 旨

大規模な時空間データの尤度関数を評価することは非常に困難であることが多い。ガウス尤度関数の計算において、高次元の逆行列および行列式の演算に時間がかかるためである。そこで、時系列の尤度関数の近似に使われる Whittle 尤度関数を不規則に位置する時空間データに一般化した、一般化 Whittle 尤度関数を提案する。この近似尤度関数は割り算のみの演算でガウス尤度関数を非常に高速に計算することができる。一般化 Whittle 尤度関数を最大化することで得られるパラメータ推定量の一致性、漸近正規性を示し、実際の推定パフォーマンスをシミュレーションおよび関東地方の地価データを用いて検証する。

キーワード：スペクトル密度関数，ピリオドグラム，離散フーリエ変換，グラム・シュミット直交化，対数尤度関数，定常空間過程。

## 1. はじめに

時空間データとは時間と場所の変化に応じて観測されるデータのことで、主に2つに分類される。ひとつが不規則な時地点において観測されるデータで point referenced data とよばれる。多数の観測地点で観測された地価や大気汚染物質の値などが代表的な例である。もう一つが適当な区分にしたがって集計して離散データ化されたもので mesh data とよばれる。県ごとに集計されたある疾病の発病率などが典型的である。

本論考では前者の point referenced data の分析法を扱う。時空間上のランダムな地点  $s_1, \dots, s_n \in R^d$  において観測された  $Y(s_1), \dots, Y(s_n)$  に対し、次の時空間回帰モデル

$$(1.1) \quad Y(s_j) = b_0 + b_1\alpha_1(s_j) + \dots + b_m\alpha_m(s_j) + X(s_j) + Z_j, \quad j = 1, \dots, n,$$

に従うとする。ここで  $X(s), s \in R^d$  は未観測の定常過程、 $Z_j$  は平均 0、分散  $\sigma_0^2$  の独立な誤差項、 $\alpha_1(s), \dots, \alpha_m(s)$  は地点  $s$  における説明変数列、 $b_0, \dots, b_m$  は回帰係数である。本論考では、定常過程  $X(s)$  の相関構造を時空間データ  $Y(s_1), \dots, Y(s_n)$  から推定することを目的とする。なお表現を見やすくするため観測地点  $s_j, j = 1, \dots, n$  を平面上  $R^2$  の点として記述することにするが、以下の議論は任意の次元に一般化される。

標本サイズが比較的少ない場合では正規尤度関数による最尤法による相関構造の分析が標準的である (Mardia and Marshall, 1984)。しかし正規尤度関数は逆行列と行列式の評価を含むため、標本サイズが大きい大規模な時空間データの最尤法によるパラメータ推定は時間がかかり

<sup>†</sup> 東北大学大学院 経済学研究科：〒980-8576 宮城県仙台市青葉区川内

現実的には難しい。したがって、大規模な時空間データの尤度近似法は近年の大きな関心をよぶテーマとなっている。Stein et al. (2004)は大規模な時空間データの尤度関数の近似法を提案している。これは、時系列解析で使われる自己回帰モデルの考え方を空間過程にあてはめたものである。アイデアがわかりやすく近似精度も良いので、大規模な時空間データの尤度計算法としては標準的なものと考えることができる。その他の尤度近似法として、Kaufman et al. (2008)は Covariance tapering とよばれる近似法を提案している。これは、時系列解析で使われる移動平均モデルの考え方を適用したもので、ある一定の距離を超える時地点間の相関を無視し0とみなして尤度近似を行う。これらの方法はいずれも空間領域において尤度関数を近似する方法と考えることができる。

本論考では空間領域に代えて、周波数領域における尤度近似法を提案する。Matsuda and Yajima (2009)は時系列解析における Whittle 尤度関数を不等間隔時空間データに拡張したものを提案した。ここではこれに改良を加えた新たな尤度関数を提案する。Matsuda and Yajima (2009)で提案した Whittle 尤度関数は、フーリエ展開を用いて時空間データを周波数領域に変換する。このとき、不等間隔データにおいてはフーリエ級数が直交関数系にならないため、周波数領域に変換する際に情報損失が生じて近似精度が落ちてしまう。そこで、フーリエ級数をそのまま使うのではなく、グラム・シュミット直交化法を用いてフーリエ関数系を直交化し、離散フーリエ変換による情報損失を最小化してから、Whittle 尤度関数を計算する方法を導く。

第2節ではよく知られている定常過程のスペクトル表現を紹介する。これをもとに第3節で一般化 Whittle 尤度関数を導出する過程を説明する。第4節では一般化 Whittle 法による推定量の漸近理論を証明し、第5, 6節ではそれぞれ関東地方の地価データ、シミュレーションデータを使った応用を紹介し、現実の推定パフォーマンスを明らかにする。ただし、第4節では、本論考で提案する推定量そのものの漸近理論を導くことには成功しておらず、積分近似を施した推定量の漸近理論を扱うことに注意を要する。

## 2. 定常過程のスペクトル表現

一般化 Whittle 法を導入するために、まず定常過程のスペクトル表現を紹介する。2次元空間上の平均0の定常過程  $X(s)$ ,  $s \in R^2$  の共分散関数  $\gamma(u) = EX(s)X(s-u)$ ,  $u \in R^2$  に対し、

$$(2.1) \quad \gamma(u) = \int_{R^2} f(\lambda) e^{iu'\lambda} d\lambda$$

を満たすスペクトル密度関数  $f(\lambda)$ ,  $\lambda \in R^2$  が存在することを仮定する。このとき、適当な仮定の下で  $X(s)$  は次のスペクトル表現を持つことが知られている (Stein, 1999, p. 23)。

$$(2.2) \quad X(s) = \int_{R^2} \exp(i\lambda's) M(d\lambda),$$

ここで  $M$  はランダム測度とよばれるもので、任意のボレル集合  $\Delta \in R^2$  に対し、

$$EM(\Delta) = 0,$$

$$EM|\Delta|^2 = \int_{\Delta} f(\lambda) d\lambda,$$

を満たし、さらに任意の排反なボレル集合  $\Delta_1, \Delta_2$  に対し

$$EM(\Delta_1)M(\Delta_2) = 0,$$

を満たす。特に  $X(s)$  が実数値のとき、ランダム測度は  $M(d\lambda) = \overline{M(-d\lambda)}$  を満たすことが必要

で、スペクトル表現は

$$(2.3) \quad X(s) = 2 \int_{R_2^+} \cos(i\lambda' s) M_1(d\lambda) + 2 \int_{R_2^+} \sin(\lambda' s) M_2(d\lambda),$$

に帰着する. ここで  $R_2^+$  は  $\{(x, y) \in R_2 | y > 0\}$  であり,  $M_1, M_2$  はそれぞれランダム測度  $M$  の実数部および虚数部である. さらに  $M_1, M_2$  は互いに無相関な測度で  $E|M_1(\Delta)|^2 = E|M_2(\Delta)|^2 = 1/2 \int_{\Delta} f(\lambda) d\lambda$  を満たす (Brockwell and Davis, 1991, p. 163).

スペクトル表現は, 定常過程が様々な周波数の波の和で与えられ, その振幅が互いに無相関な確率変数となることを示している. 数学的には興味深い表現であるが, このままでは統計分析の道具としては使いにくい. 積分をリーマン近似してスペクトル表現を回帰モデルの形式に書き直す.

スペクトル表現(2.3)の積分区間  $R_2^+$  を有限域  $D \subset R_2^+$  で近似し, さらに  $D$  を  $\omega_j, j=1, \dots, p$  を中心とする長方形の集合  $\{\delta_j, j=1, \dots, p\}$  に分割する. 被積分関数を  $\delta_j$  上で定数となるような階段関数で近似すると, スペクトル表現のリーマン近似

$$(2.4) \quad \tilde{X}(s) = \sum_{j=1}^p \phi_j \cos(\omega_j' s) + \psi_j \sin(\omega_j' s),$$

を得る. ここで係数は  $\phi_j = 2M_1(\delta_j), \psi_j = 2M_2(\delta_j)$  であり, 互いに無相関で

$$(2.5) \quad \begin{aligned} E\phi_j &= E\psi_j = 0, \\ E|\phi_j|^2 &= E|\psi_j|^2 = 2 \int_{\delta_j} f(\lambda) d\lambda \simeq 2|\delta_j|f(\omega_j) \end{aligned}$$

を満たす. リーマン近似表現  $\tilde{X}(s)$  の近似誤差  $X(s) - \tilde{X}(s)$  を  $\varepsilon(s)$  とおくと, 空間定常過程を次の回帰モデルの形式

$$(2.6) \quad X(s) = \sum_{j=1}^p \{\phi_j \cos(\omega_j' s) + \psi_j \sin(\omega_j' s)\} + \varepsilon(s)$$

で表現することができる. ここで  $\varepsilon(s)$  はスペクトル表現(2.3)において  $D$  の外側の積分値とリーマン近似誤差の和となる.

### 3. 一般化 Whittle 尤度関数

空間回帰モデル(1.1)における一般化 Whittle 尤度関数を導出する. ただし, 観測地点  $s_1, \dots, s_n$  は長方形区間  $[0, A_1] \times [0, A_2]$  のなかでランダムに位置しているものとする.  $X(s)$  のスペクトル密度関数がパラメータ  $\theta = (\theta_1, \dots, \theta_q)'$  によって  $f(\lambda; \theta)$  とモデル化されているとき, パラメータ  $\theta$  を推定するための一般化 Whittle 尤度関数を定義する.

スペクトル表現のリーマン近似モデル(2.6)を(1.1)に代入する. ただし, リーマン近似に使う周波数をフーリエ周波数

$$\omega_j = \left( \frac{2\pi j_1}{A_1}, \frac{2\pi j_2}{A_2} \right)$$

に設定し,  $(j_1, j_2)$  は  $R_2^+$  上の格子点でフーリエ周波数が  $D$  に含まれる範囲を動くものとし, これらの格子点によるフーリエ周波数を原点からの距離が近い順番に並べなおして,  $\omega_1, \dots, \omega_p$  とおくことにする.  $X(s)$  のスペクトル表現を代入した回帰モデルに対し, 従属変数を  $y = (Y(s_1), \dots, Y(s_n))'$ , 独立変数を  $\alpha_j = (\alpha_j(s_1), \dots, \alpha_j(s_n))'$ ,  $j = 1, \dots, m$ ,  $c_j = (\cos(\omega_j s_1), \dots,$

$\cos(\omega_j s_n))'$ ,  $s_j = (\sin(\omega_j s_1), \dots, \sin(\omega_j s_n))'$ ,  $j = 1, \dots, p$ , 回帰係数を  $b = (b_0, \dots, b_m)$ ,  $\Phi = (\phi_1, \psi_1, \dots, \phi_p, \psi_p)'$ , 誤差項を  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)'$ ,  $Z = (Z_1, \dots, Z_n)'$  とおく. さらに計画行列を  $B = (1, \alpha_1, \dots, \alpha_m)$ ,  $F = (c_1, s_1, \dots, c_p, s_p)$  とおくと, 回帰モデル(1.1)は行列表記

$$(3.1) \quad y = Bb + F\Phi + \varepsilon + Z$$

をもつ.

$B$  の列ベクトルと  $F$  の列ベクトルを合わせて  $a_1, \dots, a_{m+2p+1}$  とおき, 左から順にグラム・シュミットの直交化法によって正規直交基底  $e_1, \dots, e_{m+2p+1}$  に変換する. すなわち,

$$(3.2) \quad \begin{aligned} e_1 &= \frac{a_1}{\|a_1\|}, \\ e_i &= \frac{a_i - \sum_{j=1}^{i-1} (a_i' e_j) e_j}{\|a_i - \sum_{j=1}^{i-1} (a_i' e_j) e_j\|}, \quad i = 2, \dots, m + 2p + 1, \end{aligned}$$

によって直交化する. ここで  $\|\cdot\|$  は通常のユークリッドノルムである. 最初の  $m+1$  個の正規直交基底をのぞいて  $K = (e_{m+2}, \dots, e_{m+2p+1})$  とおき,  $K'$  を左から(3.1)にかけると右辺第一項が消えて,

$$(3.3) \quad K'y = K'F\Phi + K'(\varepsilon + Z),$$

を得る. そして

$$\Psi = K'y$$

を有限フーリエ変換と定義することにする. ここで  $\varepsilon$  は, スペクトル表現(2.2)を(2.4)で近似したときの近似誤差であるが, これを平均0, 分散

$$\tau(\theta) = \int_{D^c} f(\lambda; \theta) d\lambda$$

の独立系列で近似することになると, (3.3)より有限フーリエ変換  $\Psi = K'y$  の正規対数尤度関数は,

$$(3.4) \quad \log L(\theta, \sigma_0^2) = -\frac{1}{2} \Psi' V(\theta)^{-1} \Psi - \frac{1}{2} \log \|V(\theta)\|$$

となる. ここで

$$V(\theta) = K'FG(\theta)F'K + (\sigma_0^2 + \tau(\theta))I_{2p},$$

さらに,  $G(\theta)$  は  $\Phi$  の分散で, (2.5)よりこれを

$$\frac{8\pi^2}{A_1 A_2} \text{diag}(f(\omega_1; \theta), f(\omega_1; \theta), \dots, f(\omega_p; \theta), f(\omega_p; \theta))$$

によって近似している.

しかしながら, (3.4)を計算することは  $p$  が大きい場合, 難しい.  $K'F$  が上三角行列になって  $V(\theta)$  が対角行列にならないためである. そこで, 本論考では  $K'F$  をその対角成分をそのままに非対角成分を0に置き換えた行列

$$Q = \text{diagonal}(K'F)$$

で近似することを提案する. そのとき,  $V(\theta)$  は対角行列になって

$$(3.5) \quad I_j = \frac{A_1 A_2}{16\pi^2} \frac{\Psi_j^2}{Q_{jj}^2},$$

$$\kappa_j(\theta) = \frac{A_1 A_2 \sigma_0^2 + \tau(\theta)}{16\pi^2 Q_{jj}^2}, \quad j = 1, \dots, 2p,$$

とおくと, (3.4) は, パラメータ  $\theta$  に依存しない定数項を除いて,

$$(3.6) \quad -\frac{1}{2} \sum_{j=1}^p \left[ \frac{I_{2j-1}}{\frac{1}{2}f(\omega_j; \theta) + \kappa_{2j-1}(\theta)} + \frac{I_{2j}}{\frac{1}{2}f(\omega_j; \theta) + \kappa_{2j}(\theta)} \right. \\ \left. + \log \left\{ \frac{1}{2}f(\omega_j; \theta) + \kappa_{2j-1}(\theta) \right\} + \log \left\{ \frac{1}{2}f(\omega_j; \theta) + \kappa_{2j}(\theta) \right\} \right]$$

に帰着する. これを一般化 Whittle 尤度関数と定義する. メッシュデータの場合のアナロジーにより  $I_{2j-1} + I_{2j}$  を  $\omega_j$  におけるピリオドグラム  $I(\omega_j)$  と定義する. なお, メッシュデータに対しては,  $Q_{jj}^2 = n/2$  となるため, 一般化 Whittle 尤度関数は既存の Whittle 尤度関数に一致する. 一般化 Whittle 尤度関数はデータ数が多い場合に有効かつ高速な近似尤度計算法を与えることが期待される.

ここで Matsuda and Yajima (2009, p. 195, Definition 2) によるピリオドグラムの定義:

$$(3.7) \quad I_M(\lambda) = |J(\lambda)|^2, \\ J(\lambda) = (2\pi)^{-1} \sqrt{A_1 A_2} n^{-1} \sum_{k=1}^n X(s_k) \exp(-\lambda' s_k),$$

および本稿が提案するピリオドグラムとの差異についてふれておく. 前者のピリオドグラム  $I_M(\omega_j)$  は, (2.6) の回帰モデルにおいて独立変数

$$f_{2j-1} = (\cos(\omega_j s_1), \dots, \cos(\omega_j s_n))', \\ f_{2j} = (\sin(\omega_j s_1), \dots, \sin(\omega_j s_n))',$$

それぞれに単回帰してその最小二乗推定値を二乗して足したものを, 後者は  $f_1, \dots, f_{2p}$  を直交基底  $e_1, \dots, e_{2p}$  に変換してから  $e_{2j-1}, e_{2j}$  に単回帰してその最小二乗推定値を二乗して足したものに相当する. したがって, メッシュデータの場合には,  $f_1, \dots, f_{2p}$  は直交しているの両者の定義は一致する. 一方, 不等間隔データの場合には前者の回帰係数は一致性を失うが, 後者は直交化によってその情報損失が抑えられる. つまり不等間隔データにおいて本論考で提案する有限フーリエ変換は (3.7) に比べて, データの情報損失が少なくなるように周波数領域に変換していることになる.

最後に  $p$  の決定についてふれる.  $n \times 2p$  行列  $K$  の列ベクトルは直交させなければならないため,  $2p + m + 1 \leq n$  の範囲で選ぶ. ただし,  $2p + m + 1$  が  $n$  に近づくと  $a_1, \dots, a_{m+2p+1}$  の一次独立性が保証されず, グラム・シュミット直交化の手続きが不安定になってしまう. そこで  $2p + m + 1$  が  $n$  に比べてやや小さくなるように  $p$  をとり, 直交化法が問題なく働く範囲で決めるとよい.

#### 4. 漸近理論

本節では  $m=0$  の場合に一般化 Whittle 尤度関数 (3.6) を最大化する推定量  $\hat{\theta}$  の一貫性, 漸近正規性を証明する.  $m=0$  の場合に限定する理由は (3.3) において  $K'F \simeq \sqrt{n/2} I_{2p}$  の近似が正当化されてピリオドグラムが単純な形に帰着し, Matsuda and Yajima (2009) の証明法が適用可能となるからである. ただし,  $m=0$  を仮定してもこのままの推定量の漸近理論を構築することができない. テイパーとよばれるものを用いて有限フーリエ変換を調整する必要がある.

$h(x_1, x_2)$  を  $[0, 1] \times [0, 1]$  の外側では 0 となる関数とし,  $H = \int_{[0,1] \times [0,1]} |h(x_1, x_2)|^2 dx_1 dx_2$  として

$$\tilde{\Psi}_j = H^{-1/2} \sum_{i=1}^n K_{ij} y_i h\left(\frac{s_{i,1}}{A_1}, \frac{s_{i,2}}{A_2}\right), \quad j=1, \dots, 2p$$

を有限フーリエ変換とする. テイパーが定数ならば元の定義に一致する. ピリオドグラム, 一般化 Whittle 尤度の定義はこの有限フーリエ変換を使って前節と同様に与えられる. 一般化 Whittle 尤度関数 (3.6) をテイパーで調整してもなお漸近理論を構築することには残念ながら成功していない. 漸近理論を構築するためには, (3.6) を積分近似した尤度関数

$$(4.1) \quad -\frac{1}{2} \int_D \left[ \frac{I_1(\lambda)}{\frac{1}{2}f(\lambda; \theta) + \kappa_1(\theta, \lambda)} + \frac{I_2(\lambda)}{\frac{1}{2}f(\lambda; \theta) + \kappa_2(\theta, \lambda)} + \log \left\{ \frac{1}{2}f(\lambda; \theta) + \kappa_1(\theta, \lambda) \right\} + \log \left\{ \frac{1}{2}f(\lambda; \theta) + \kappa_2(\theta, \lambda) \right\} \right] d\lambda$$

を最大化する推定量を対象にしなければならない. ただし,  $I_1(\lambda)$ ,  $I_2(\lambda)$ ,  $\kappa_1(\theta, \lambda)$ ,  $\kappa_2(\theta, \lambda)$  は,  $\lambda$  に最も近いフーリエ周波数  $\omega_j$  に対し, それぞれ (3.5) における  $I_{2j-1}$ ,  $I_{2j}$ ,  $\kappa_{2j-1}(\theta)$ ,  $\kappa_{2j}(\theta)$  を形式的に拡張したものである. ここで,  $I_{2j-1}$ ,  $I_{2j}$  に必要な  $\Psi_{2j-1}$ ,  $\Psi_{2j}$  を計算する際, (3.2) において分子第一項だけをそれぞれ  $(\cos(\lambda s_1), \dots, \cos(\lambda s_n))'$ ,  $(\sin(\lambda s_1), \dots, \sin(\lambda s_n))'$  に置き換えなければならない. なお, (4.1) に対してしか漸近理論を構築することに成功していないが, 現実には (3.6) を最尤推定に用いることを本稿では提案する.

漸近理論を構築する前に, 漸近の意味を明らかにしておく必要がある. 時系列の場合には標本サイズの増大がそのまま漸近理論につながるが, 空間過程においては, 標本サイズとともに観測地域の広がり方も考慮に入れる必要がある. 空間過程における漸近理論にはいくつかの方法があるが, 本論考では mixed asymptotics とよばれる漸近法を用いる (仮定 (1)). 以下に漸近理論の構築に必要な仮定 (1)–(4) をあげる.

- (1)  $s_1, \dots, s_n$  を  $R^2$  平面上の観測地点とし, 各点は

$$s_j = (A_1 u_{j,1}, A_2 u_{j,2})$$

と表されるものとする. ここで  $u_j = (u_{j,1}, u_{j,2})$ ,  $j=1, \dots, n$  は独立に  $[0, 1] \times [0, 1]$  上に一様分布する確率変数である. さらに  $n$ ,  $A_1$ ,  $A_2$  は  $k$  の関数で,  $k \rightarrow \infty$  のとき,  $n = n_k \rightarrow \infty$ ,  $A_1 = A_1(k) \rightarrow \infty$ ,  $A_2 = A_2(k) \rightarrow \infty$  かつ  $A_1(k)/A_2(k) = O(1)$ ,  $\{A_1(k)A_2(k)\}^3/n_k \rightarrow 0$  を満たす.

- (2)  $\{X_t, t \in R^2\}$  はスペクトル密度関数  $f(\lambda)$ ,  $\lambda = (\lambda_1, \lambda_2) \in R^2$  をもつ平均 0 の正規定常過程であり,  $f(\lambda)$  は有界かつ適当な定数  $C$  に対し,

$$\left| \frac{\partial^2 f}{\partial \lambda_p \partial \lambda_q}(\lambda) \right| \leq C, \quad p, q=1, 2,$$

が成り立つ.

- (3) テイパー  $h(x_1, x_2)$  は  $[0, 1] \times [0, 1]$  内にサポートをもち, 1 階偏微分可能かつ  $0 \leq m_1$ ,  $m_2 \leq 2$  に対し,

$$\frac{\partial^{m_1+m_2} h(x_1, x_2)}{\partial x_1^{m_1} \partial x_2^{m_2}}$$

が存在する.

- (4) スペクトル密度関数はパラメータ  $\theta \in \Theta$  によってモデル化され,  $\Theta$  の内点  $\theta_0$  を真の値として  $f(\theta_0, \lambda) = f(\lambda)$  が成り立つ.  $f(\lambda; \theta)$  は  $\Theta \times D$  上で正值をとり,  $\lambda \in D$  に対し  $\theta$  に

関して二回偏微分可能である. さらに  $\theta_1 \neq \theta_2$  ならば  $f(\theta_1, \lambda) \neq f(\theta_2, \lambda)$  となる  $\lambda \in D$  の集合は正のルベグ測度を持つ.

**定理 1.** (1.1)において,  $m=0$  とする. 仮定 1-4 の下で  $k \rightarrow \infty$  のとき,  $\hat{\theta}_k$  は  $\theta_0$  に確率収束し, かつ漸近正規性

$$(A_1 A_2)^{1/2} (\hat{\theta}_k - \theta_0) \rightarrow N(0, 2b\Gamma(\theta_0)^{-1})$$

が成立する. ここで

$$b = \left\{ \int_{[0,1]^2} |h(x)|^4 dx \right\} \left\{ \int_{[0,1]^2} |h(x)|^2 dx \right\}^{-2},$$

$$\Gamma(\theta) = \frac{1}{2\pi^2} \int_D \left( \frac{\partial \log f(\lambda; \theta)}{\partial \theta} \right) \left( \frac{\partial \log f(\lambda; \theta)}{\partial \theta} \right)' d\lambda.$$

以下, 証明に必要な補題を述べた後, 本定理の証明を行う. なお,  $a = (a_1, \dots, a_n)'$ ,  $b = (b_1, \dots, b_n)'$  に対し,  $\langle a, b \rangle$  は内積  $\sum_{i=1}^n a_i b_i$  を表すものとする.

**補題 1.**  $\{c_1, s_1, \dots, c_p, s_p\} = \{f_1, \dots, f_{2p}\}$  とおく.  $n \rightarrow \infty$  のとき,

- (1)  $\frac{2}{n} \langle f_i, f_i \rangle - 1 = O_p(n^{-1/2})$ ,  $i = 1, \dots, 2p$ ,
- (2)  $\frac{1}{n} \langle f_i, f_j \rangle = O_p(n^{-1/2})$ ,  $i \neq j$ .

**証明.**  $f_i = (\cos(\omega'_i s_1), \dots, \cos(\omega'_i s_n))'$  のときを例として証明する.

$$\langle f_i, f_i \rangle = \sum_{k=1}^n \cos^2(\omega_i s_k) = \frac{n}{2} + \frac{1}{2} \sum_{k=1}^n \cos(2\omega_i s_k).$$

したがって,  $\omega_i$  がフーリエ周波数であることにより,

$$E \left( \frac{2}{n} \langle f_i, f_i \rangle - 1 \right) = \int_{[0, A_1] \times [0, A_2]} \cos(2\omega'_i u) A_1^{-1} A_2^{-1} du = 0$$

となり,

$$\text{Var} \left( \frac{2}{n} \langle f_i, f_i \rangle - 1 \right) = n^{-1} \int_{[0, A_1] \times [0, A_2]} \cos^2(2\omega'_i u) A_1^{-1} A_2^{-1} du = O(n^{-1})$$

であることから証明される. (2) も同様.

**補題 2.**  $\hat{f}_j$  を  $f_j$  の  $f_1, \dots, f_{j-1}$  上への射影とする.

$$\frac{\langle \hat{f}_j, \hat{f}_j \rangle}{\langle f_j, f_j \rangle} = O_p(n^{-1} A_1 A_2), \quad j = 1, \dots, 2p.$$

**証明.**

$$\frac{\langle \hat{f}_j, \hat{f}_j \rangle}{\langle f_j, f_j \rangle} \leq \sum_{k=1}^{j-1} \frac{\langle f_k, f_j \rangle^2}{\langle f_k, f_k \rangle \langle f_j, f_j \rangle}, \quad j = 1, \dots, 2p.$$

有限区間  $D$  に含まれるフーリエ周波数の数は高々  $A_1 A_2$  のオーダーであることと補題 1 を適用することで証明される.

**補題 3.**  $I_M(\lambda)$  を (3.7) で示した Matsuda and Yajima (2009, Definition 2) によるピリオドグラムであるとする. (3.5) 式で定義するピリオドグラム  $I(\omega_j) = I_1(\omega_j) + I_2(\omega_j)$  を任意の  $\lambda$  に (4.1) のように形式的に拡張した  $I(\lambda)$  に対し, 次の結果が成立する.

$$|I(\lambda) - I_M(\lambda)| = o_p((A_1 A_2)^{-1/2}).$$

**証明.** まず,

$$(4.2) \quad \frac{\Psi_j}{Q_{jj}} - \frac{2\langle y, f_j \rangle}{n} = o_p((A_1 A_2)^{-1}), \quad j=1, \dots, 2p,$$

を示す.  $\tilde{f}_j$  を  $f_j$  の  $f_1, \dots, f_{j-1}$  上への射影残差とする. このとき,

$$\frac{\Psi_j}{Q_{jj}} = \frac{\langle y, e_j \rangle}{\langle e_j, f_j \rangle} = \frac{\langle y, \tilde{f}_j \rangle}{\langle \tilde{f}_j, \tilde{f}_j \rangle} = \frac{\langle y, \tilde{f}_j \rangle}{\langle \tilde{f}_j, \tilde{f}_j \rangle}$$

であるから,

$$\begin{aligned} \left| \frac{\Psi_j}{Q_{jj}} - \frac{2\langle y, f_j \rangle}{n} \right| &\leq \left| \frac{\langle y, \tilde{f}_j \rangle}{\langle \tilde{f}_j, \tilde{f}_j \rangle} - \frac{\langle y, f_j \rangle}{\langle f_j, f_j \rangle} \right| + \left| \frac{\langle y, f_j \rangle}{\langle f_j, f_j \rangle} - \frac{2\langle y, f_j \rangle}{n} \right| \\ &+ \left| \frac{\langle y, f_j \rangle}{\langle f_j, f_j \rangle} - \frac{2\langle y, f_j \rangle}{n} \right| \\ &= B_1 + B_2 + B_3 \end{aligned}$$

とおく.

$$B_1 \leq \frac{\sqrt{\langle y, y \rangle} \sqrt{\langle \tilde{f}_j - f_j, \tilde{f}_j - f_j \rangle}}{\langle \tilde{f}_j, \tilde{f}_j \rangle} = \sqrt{\frac{\langle y, y \rangle}{\langle \tilde{f}_j, \tilde{f}_j \rangle}} \sqrt{\frac{\langle \tilde{f}_j - f_j, \tilde{f}_j - f_j \rangle}{\langle \tilde{f}_j, \tilde{f}_j \rangle}} = O_p(n^{-1/2}(A_1 A_2)^{1/2}),$$

$$B_2 = \left| \frac{\langle y, f_j \rangle}{\langle f_j, f_j \rangle} \right| \left| 1 - \frac{\langle \tilde{f}_j, \tilde{f}_j \rangle}{\langle f_j, f_j \rangle} \right| = \left| \frac{\langle y, f_j \rangle}{\langle f_j, f_j \rangle} \right| \left| \frac{\langle \tilde{f}_j, \tilde{f}_j \rangle}{\langle f_j, f_j \rangle} \right| = O_p(n^{-1}(A_1 A_2)^{1/2}),$$

$$B_3 = \left| \frac{\langle y, f_j \rangle}{\langle f_j, f_j \rangle} \right| \left| 1 - \frac{2}{n} \langle f_j, f_j \rangle \right| = O_p(n^{-1/2}(A_1 A_2)^{-1/2}).$$

仮定(1)より,  $B_1, B_2, B_3$  は高々  $o_p((A_1 A_2)^{-1})$  であることがわかる.

(4.2) より,

$$\begin{aligned} |I(\omega_j) - I_M(\omega_j)| &\leq \frac{A_1 A_2}{16\pi^2} \left| \frac{\Psi_{2j-1}^2}{Q_{2j-1, 2j-1}^2} - \frac{4\langle y, f_{2j-1} \rangle^2}{n^2} \right| + \frac{A_1 A_2}{16\pi^2} \left| \frac{\Psi_{2j}^2}{Q_{2j, 2j}^2} - \frac{4\langle y, f_{2j} \rangle^2}{n^2} \right| \\ &= \frac{\sqrt{A_1 A_2}}{16\pi^2} \left| \frac{\Psi_{2j-1}}{Q_{2j-1, 2j-1}} - \frac{2\langle y, f_{2j-1} \rangle}{n} \right| \sqrt{A_1 A_2} \left| \frac{\Psi_{2j-1}}{Q_{2j-1, 2j-1}} + \frac{2\langle y, f_{2j-1} \rangle}{n} \right| \\ &+ \frac{\sqrt{A_1 A_2}}{16\pi^2} \left| \frac{\Psi_{2j}}{Q_{2j, 2j}} - \frac{2\langle y, f_{2j} \rangle}{n} \right| \sqrt{A_1 A_2} \left| \frac{\Psi_{2j}}{Q_{2j, 2j}} + \frac{2\langle y, f_{2j} \rangle}{n} \right| \\ &= o_p((A_1 A_2)^{-1/2}) \end{aligned}$$

となる. これはフーリエ周波数上で評価されたものであるが, フーリエ周波数以外にも自然に拡張される.

**定理 1 の証明.** 補題 3 より Lemma 7 (Matsuda and Yajima, 2009) は, ピリオドグラムを本稿が定義したピリオドグラムに置き換えても成立する. 一致性は Theorem 1 (Matsuda and Yajima, 2009) と同様にして, 漸近正規性は Theorem 2 (Matsuda and Yajima, 2009) と同様に Lemma 7 を使って証明される.

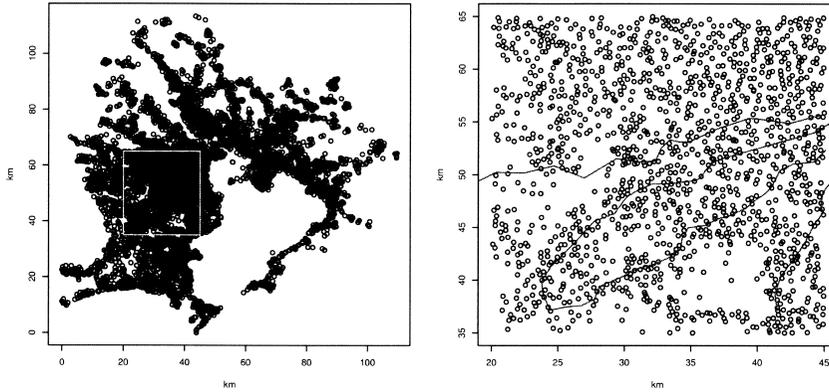


図 1. 関東地区公示地価データの 5573 標本地点と南西部地域 (左図長方形部分) 1431 標本地点の拡大図. 右図の曲線は電車の路線を示す.

### 5. 地価データへの応用

本節では本稿で提案した一般化 Whittle 尤度 (3.6) を地価データの分析に応用する. 地価データは 2001 年に国土交通省によって収集された公示地価の一部で, 5573 標本地点の平米単価とその緯度と経度とともに都心からの時間距離と最寄駅からの距離とともに記録されている. 図 1 の左図は標本地点を単位を km とする通常の直角座標に直した散布図である. 全データを定常モデルを使って分析対象とするには大きすぎるので, 一部分を抜き出して分析を行う. 図 1 の右図で表されている 1431 地点の地価データを対象に分析をすすめる.

地価データは場所の特性に依存した構造をもつので一定の平均値をもつ定常過程の実現値とみなすことはできない. そこで空間回帰モデル (1.1) を本地価データにあてはめることにする.  $\alpha_1(s), \alpha_2(s)$  をそれぞれ地点  $s$  における都心からの距離, 最寄駅からの距離とし, 地価  $Y(s)$  を回帰モデル

$$(5.1) \quad Y(s_j) = c_0 + c_1 b_1(s_j) + c_2 b_2(s_j) + X(s_j) + Z_j, \quad j = 1, \dots, 1431,$$

で表す. ここで  $X(s)$  は定常過程で, そのスペクトル密度関数は Matérn class (Gelfand et al., 2010, Sec. 2.7 参照)

$$f(\lambda) = \frac{\phi}{(\alpha^2 + \|\lambda\|^2)^{\nu+1}}, \quad \lambda \in R_2$$

でモデル化されているとする. Matérn class の自己共分散関数はこのスペクトル密度の逆フーリエ変換によって得られ

$$(5.2) \quad \gamma(h) = \frac{\sigma_1}{2^{\nu-1}\Gamma(\nu)} \left( \frac{2\sqrt{\nu}\|h\|}{\rho} \right)^\nu K_\nu \left( \frac{2\sqrt{\nu}\|h\|}{\rho} \right), \quad h \in R_2,$$

となることが知られている. 自己共分散を記述するパラメータ  $(\nu, \sigma_1, \rho)$  はそれぞれ平滑化パラメータ, シル, レンジとよばれ, スペクトル密度関数を記述するパラメータ  $(\nu, \alpha, \phi)$  と次の関係で結ばれている.

$$\alpha = \frac{2\sqrt{\nu}}{\rho}, \quad \phi = \frac{\nu\alpha^{2\nu}\sigma_1}{\pi}.$$

このモデルに対し, 本論考で提案した一般化 Whittle 尤度関数 (3.6) を最大化することでパラメータの推定を行う. 比較のため Stein et al. (2004) により提案された推定法 (以下 Stein 法)

表 1. 地価データに対する一般化 Whittle 法(Matsuda), Stein et al.(2004)の方法, Matsuda and Yajima (2009) の方法による最尤推定値 (MLE) とその標準誤差 (SE).

		$\nu$	$\sigma_1$	$\rho$	$\sigma_0$	Prediction RMSE
Matsuda	MLE	0.48	34.2	4.94	4.29	
	SE	(0.17)	(7.77)	(1.55)	(2.07)	5.87
Stein	MLE	0.58	31.1	4.96	3.38	
	SE	(0.12)	(6.14)	(1.25)	(0.79)	5.84
MY	MLE	3.17	18.27	2.90	1.01	
	SE	(2.76)	(3.67)	(0.45)	(3.88)	6.33

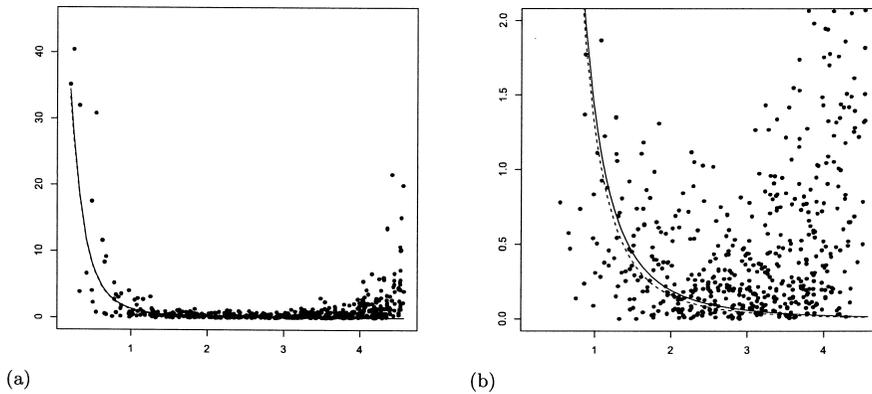


図 2. ピリオドグラムとスペクトル推定値(Matsuda(破線), Stein 法(実線)). (a)は全体図, (b)は高周波部分に拡大したものである. ここで横軸は周波数である.

および一般化 Whittle 法の改良元となった Matsuda and Yajima (2009)による推定法(以下 MY 法)による推定も同時に行う. Stein 法は, 大規模な時空間データの尤度関数を本論考の方法の周波数領域ではなく時空間領域で近似を行う代表的な近似法とみなされている. なお, 3つの推定量の推定の良さを比較するため, 1431 地点のデータから 100 地点をランダムに選び, 残りの 1331 地点を使ってパラメータの推定を行い, 推定したパラメータを使って 100 地点の予測量を構成した. パラメータ推定結果と予測誤差の二乗平均を計算した結果を表 1 に紹介している. 図 2 (a), (b)にピリオドグラムと推定されたスペクトル密度関数を示す. (a)が低周波部分, (b)が高周波部分に拡大したものである.

まず, 一般化 Whittle 法による予測誤差は MY 法のそれを改善していることから, その尤度近似の精度も向上していると考えられる. 次に一般化 Whittle 法と Stein 法の推定値はかなり似通っており, 前者の予測誤差は Stein 法のそれと比べてやや劣っている. また推定されたスペクトル密度関数を比較すると, 両者の違いはほとんど見られず, 違いは前者による推定値が低周波部分でややスペクトルの値が Stein 法の推定値に比べて高くなっていることくらいである. 一般化 Whittle 法は Stein 法とほぼ同等な推定パフォーマンスを有する尤度近似法であることを示している.

表 2. Case 1:  $\nu=0.5$ ,  $\sigma_1=30.5$ ,  $\rho=4.6$ ,  $\sigma_0=3.46$  の Matérn class で発生させたシミュレーションデータに対する一般化 Whittle 法 (Matsuda), Stein et al. (2004) の方法, Matsuda and Yajima (2009) の方法による最尤推定値のバイアス (bias) と平均二乗誤差の平方根 (RMSE) .

		$\nu$	$\sigma_1$	$\rho$	$\sigma_0$	Prediction RMSE
Matsuda	bias	-0.02	0.53	-1.30	-0.30	
	RMSE	0.12	5.82	1.57	1.55	8.89
Stein	bias	0.038	-2.18	-0.32	-0.12	
	RMSE	0.17	6.63	1.29	1.45	8.89
MY	bias	1.20	-7.08	-1.88	-2.32	
	RMSE	1.69	8.03	1.95	2.40	9.45

表 3. Case 2:  $\nu=3.0$ ,  $\sigma_1=30.5$ ,  $\rho=4.6$ ,  $\sigma_0=3.46$  の Matérn class で発生させたシミュレーションデータに対する一般化 Whittle 法 (Matsuda), Stein et al. (2004) の方法, Matsuda and Yajima (2009) の方法による最尤推定値のバイアス (bias) と平均二乗誤差の平方根 (RMSE) .

		$\nu$	$\sigma_1$	$\rho$	$\sigma_0$	Prediction RMSE
Matsuda	bias	-0.89	0.21	-0.25	0.44	
	RMSE	1.17	8.29	0.75	0.52	4.08
Stein	bias	1.39	-2.79	-0.26	0.01	
	RMSE	2.74	8.22	0.71	0.15	4.05
MY	bias	1.10	-8.68	-1.43	-2.44	
	RMSE	2.45	9.92	1.48	2.44	4.39

## 6. シミュレーション

本節は地価データ分析で実証した一般化 Whittle 法の性質をシミュレーションを使ってたしかめることを目的とする. 比較対象として, 前節と同様に Stein 法および MY 法を考える. MY 法は漸近的には一般化 Whittle 法と同等であることを第 4 節補題 3 において示したが, 有限標本において一般化 Whittle 法がその推定パフォーマンスをどの程度改善するのかを検証する.

地価データと同じ 1431 標本地点の座標をそのまま使い, 地価を Matérn class の共分散関数 (5.2) において Case 1:  $\nu=0.5$ ,  $\sigma_1=30.5$ ,  $\rho=4.6$ ,  $\sigma_0=3.46$  および Case 2:  $\nu=3$ ,  $\sigma_1=30.5$ ,  $\rho=4.6$ ,  $\sigma_0=3.46$  として発生させた定常データに対して, 一般化 Whittle 尤度 (3.6), Stein 法および MY 法を使ってパラメータ  $\nu, \sigma_1, \rho, \sigma_0$  を推定する. これを 100 回繰り返し, 推定精度をバイアスと平均二乗誤差によって比較したものが表 2, 3 である. ここで 3 推定量の良さを比較するため, ランダムに 100 地点を選んで残りの 1331 地点を使って推定を行い 100 地点の予測を行って予測誤差の比較も紹介している. さらに各方法の推定 1 回あたりの平均実行速度を示すと, 一般化 Whittle 法, Stein 法 および MY 法はそれぞれ 11sec., 65sec., 45sec. となっている. ただ, Stein 法は尤度計算に労力を使い, 一般化 Whittle 法は尤度計算の前の直交化に時間がかかる. そのため一般化 Whittle 法の尤度計算そのものは MY 法とほぼ同一の計算時間となるが, ここでは前処理の直交化を含めた計算時間を記述している. したがって, 最適化のための繰り返しが多い(少ない)ほど, また標本サイズが増える(減る)ほど一般化 Whittle 法の有効性は Stein 法に比べて強まる(弱まる)ことに留意する必要がある. なお, Dell Vostro 430, Intel Core i7 CPU, 860@2.80Gh の計算機においてプログラミング言語 Ox によって各手法を記述し, その最適化パッケージ MaxBFGS によって各尤度関数最大化を行っている.

シミュレーション結果は、地価データ分析で得られた一般化 Whittle 法の性質をほぼ追認するものとなっている。すなわち一般化 Whittle 法は MY 法のものに比べて予測誤差を改善しているが、Stein 法に比べるとやや劣っている。これは、一般化 Whittle 法の尤度近似精度は Stein 法とほぼ同等の水準まで改善されていることを示す。その理由として、有限標本における MY 法は第 3 節終わりで論じた単回帰近似によって無視できない情報損失を発生させる一方、三角関数系のグラム・シュミット直交化がその情報損失を抑えることが考えられる。一般化 Whittle 法は Stein 法に比べて遜色のない高速な尤度近似法であり、かつ周波数領域における尤度近似法である。空間領域における尤度近似の代表である Stein 法の代替として、一般化 Whittle 法は大規模な時空間データの分析が望まれている現況において有用な分析方法を提供することになると考えている。

### 参 考 文 献

- Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*, Springer-Verlag, New York.
- Gelfand, A., Diggle, P., Fuentes, M. and Guttop, P. (2010). *Handbook of Spatial Statistics*, CRC Press, Boca Raton, Florida.
- Kaufman, C., Schervish, M. and Nychka, D. (2008). Covariance tapering for likelihood based estimation in large spatial data sets, *Journal of the American Statistical Association*, **103**, 1556–1569.
- Mardia, K. V. and Marshall, R. J. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression, *Biometrika*, **71**, 135–146.
- Matsuda, Y. and Yajima, Y. (2009). Fourier analysis of irregularly spaced data on  $\mathbb{R}^d$ , *Journal of the Royal Statistical Society Series B*, **71**, 191–217.
- Stein, M. (1999). *Interpolation of Spatial Data*, Springer-Verlag, New York.
- Stein, M. L., Chi, Z. and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets, *Journal of the Royal Statistical Society Series B*, **66**, 275–296.

## Generalized Whittle Likelihood Method for Irregularly Spaced Data

Yasumasa Matsuda

Graduate School of Economics and Management, Tohoku University

It is sometimes difficult to evaluate the likelihood function of large irregularly spaced data. This paper proposes a method that approximates it on the frequency domain, which is considered to be a generalized version of Whittle likelihood for time series to that for irregularly spaced data. The generalized Whittle likelihood function provides a computer efficient algorithm for evaluating likelihood functions for large irregularly spaced data. After providing the consistency and asymptotic normality of the estimators that maximize the generalized Whittle likelihood function, we examine empirical performances of the Whittle likelihood method by using simulation and real data of land prices in Kanto area.