

条件付確率場の最大エントロピー推定法の提案

野田 茂¹・星谷 勝²

¹正会員 工博 鳥取大学助教授 工学部社会開発システム工学科 (〒680-0945 鳥取市湖山町南4-101)

²正会員 Ph. D. 武蔵工業大学教授 工学部土木工学科 (〒158-0087 東京都世田谷区玉堤1-28-1)

本論文では、エントロピーの概念を用いて条件付確率場を推定する理論を提案した上で、複数観測点のデータに基づいて未観測点での最大エントロピー推定値、条件付エントロピー、条件付平均エントロピーなどの性質を星谷・野田・稲田の理論(1997)と比較分析した。6つのタイプの確率場(正規、対数正規、指数、レイリー、ガンベル、一様分布)を検討した結果、1) 最大エントロピー推定値、条件付エントロピーは観測値に依存すること、2) 条件付平均エントロピーは観測配置のみに左右され、無条件平均エントロピーよりも大きくならないこと、3) 不偏推定・最小誤差分散に基づく最適推定値、条件付分散、推定誤差分散と同一の性質を有すること、4) 正規確率場のKrigingと最大エントロピー推定法の結果は等価であることなど、新たな知見が得られた。

Key Words: *stochastic fields, spatial estimation, Kriging, entropy, optimum estimator, estimation error, conditional probability density function, conditional simulation*

1. 序論

空間的に変動する確率場において、データの空間依存性を考慮しながら、観測値を用いてデータが取得されていない特定の場所での物理量を推定する手法の一つがKriging法である。Krigingには事前確率場の前提条件の違い、推定領域の大きさ、複数の異なる物理量の導入などにより、色々な手法が考案されている¹⁾。この方法によれば、各種の推定量の不確実性(推定誤差)を定量的に評価することができる。この手法を用いた空間分布の推定問題は、推定点が観測点と一致した場合、ノイズが存在しなければ、推定値と観測値が完全に同一値をとることを条件としているので、条件付推定問題と言える。

有限個の既知観測値を条件として観測前に表現された確率場を更新して得られた事後の確率場は条件付確率場と呼ばれる。物理量が正規分布に従う条件付正規確率場の推定理論は工学上最も基本的であり、かつ重要であるので、現在各分野で多用されている。最近では、未観測点でのサンプル実現値を知るため、モンテカルロ法に基づいた新しい条件付シミュレーション法が星谷²⁾、Shinozuka and Zhang³⁾によって考案された。この手法は、ブートストラップ法の概念を条件付シミュレーションに初めて導入したJournalの古典的な手法¹⁾と異なり、Krigingによる最適推定値に推定誤差分散を考慮した正規乱数を加えることにより、相互依存した複数地点のサンプル実現値を効率的に得ることを骨子としている。

一方、星谷・野田・稲田⁴⁾は、非正規分布に従う条件

付確率場において、未観測点における物理量の最適推定値、条件付分散、推定誤差分散を求める一般的な理論式を提案するとともに、相関特性を満たすように複数地点のサンプル値を実現するための条件付シミュレーション法を示している。その理論は、条件付正規確率場の推定手法と、非正規確率変数の標準正規確率変数への変換則を組み合わせ、条件付確率密度関数を定義することを基本としている。条件付確率場の推定理論に関してはこれまで多くの研究が進められてきたが、星谷らの提案手法は厳密性・効率性・有用性・汎用性に富んでおり、この分野の発展に多大な貢献をしている。

条件付確率場の推定理論は不偏性ならびに推定誤差分散の最小化を規範とするものである。これより、一般的な非正規確率場において、1) 最適推定値と条件付平均値が等しいこと、2) 条件付分散と推定誤差分散は異なること、3) 最適推定値と条件付分散は観測値に依存するが、推定誤差分散は観測値に独立であること、4) 条件付シミュレーションには推定誤差分散が関与しないことなどが明らかにされている⁴⁾。なお、条件付正規確率場では条件付分散と推定誤差分散が同一値をとるので、上記性質と異なることに注意しなければならない。

本論文は、このような不偏性・推定誤差分散の最小化に基づくのではなく、主としてShannon⁵⁾によって開拓された情報理論を用いて、非正規分布に従う確率場に関して条件付確率場を推定する理論を提案し、星谷・野田・稲田による推定理論⁴⁾との対比を行い、提案手法の性質について検討を行う。

情報理論とは、種々の通信における情報の発生や伝送を定量的に把握するために提起された理論である。この

理論では、同等に確からしい二つの事柄から、そのうちの一つを1回選ぶ情報の量を情報量の単位としている。この着想はFisher, Shannon, Wienerなどが同時に考えついたものである⁶⁾が、情報科学の誕生はShannonが論文を発表した1948年頃とされている。確率事象系の情報量はShannonのエントロピーと呼ばれ、それには結合エントロピー、条件付エントロピー、条件付平均エントロピー、相互情報量などがある⁷⁾。このエントロピーあるいは情報量はモデルの先験的な不確かさ、事後モデルの不確かさなどを測る尺度である。

エントロピーは情報源の不確定性(あるいは無秩序性)を示している。その値が大きいくほど、その情報源からの情報が発せられるかは予測がつきにくくなり、それだけ確実さがくずれているということになる。そこでこのエントロピーをうまく活用することによってモデルを決定する試みが最近行われるようになった。本城・工藤は、事前情報、観測、予測の過程に正規分布に関してよく知られた相互情報量を組み込み、相互情報量の大小によって観測条件を決めている⁸⁾。また、観測計画の評価基準として相互情報量以外の代表的な最適化基準をも用いて比較分析することにより、相互情報量の有用性を検討している。一方、吉田⁹⁾も同様にモデルパラメータを決定するのにこの相互情報量を用いている。文献9)では、未知パラメータを固有値解析によって基底ベクトルの展開係数で表し、その展開係数を逆解析する際、基底ベクトルの数を決めるのに相互情報量を最適化基準として採用している。これらの研究においては、事前に設定した情報に対し、観測を通して更新された事後情報の最適解がエントロピーの概念を用いることによって求められている。

文献2), 3)では条件付正規確率場を、文献8), 9)では無条件正規確率場を、文献4)では条件付正規確率場、5つの非正規分布(対数正規分布、指数分布、レイリー分布、グンベル分布、一様分布)に従う条件付確率場を対象に分析を実施している。正規分布のエントロピーは文献8), 9)のようによく利用され、成書7)にもわかりやすくまとめられている。しかしながら非正規分布のエントロピーについてはこれまで体系的に示されていないので、本論文では条件付非正規確率場に特有のエントロピー推定理論を新たに定式化する。

2.では問題の設定を、3.ではエントロピーを用いて条件付非正規確率場を推定するのに必要な条件付確率密度関数などを前もって定義し、4.ではエントロピーの基礎的な概念をとりまとめる。5.では非正規確率場の最大エントロピー推定量(あるいは最大エントロピー推定値)、無条件確率場の平均エントロピー、条件付確率場の条件付エントロピー、条件付平均エントロピーに関する一般式を提案して、星谷・野田・稲田の推定理

論との類似点を述べた上で、6.では提案手法を代表的な確率分布(正規分布、対数正規分布、指数分布、レイリー分布、グンベル分布、一様分布)に適用して条件付確率場の推定に必要な諸式を誘導する。7.では提案手法を用いて数値解析を実施し、最大エントロピー推定法の有用性を示す。8.で結論とする。

2. 問題の設定

ここでは正規分布ならびに非正規分布に従う非均一確率場において、離散空間位置 y_i における物理量の確率変数を x_i で表わす。この x_i については事前の試験や調査または類似環境条件における過去の経験に基づき、確率的情報が得られるとする。

今、 n 地点の確率変数 $x_i(i=1\sim n)$ よりなる確率変数ベクトル $X_n = \{x_1, x_2, \dots, x_n\}^T$ に関する事前情報として、以下の仮定を設ける。

- 1) 確率変数 x_i の周辺確率密度関数 $f_{X_i}(x_i)$ は事前に与えられる。従って観測値を得る前の任意地点 y_i における無条件平均値 μ_i 、無条件分散 σ_i^2 は既知である。
- 2) 任意の2地点 y_i と y_j における確率変数 x_i, x_j の相関係数 ρ_{ij} はモデル化された共分散 $C_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)]$ 、分散 σ_i^2, σ_j^2 によって推定されている。これより X_n の相関係数行列 R_n は $\rho_{ij}(i, j=1\sim n)$ を用いると求められる。

本論文の定式化では非均一確率場を対象にしているので、一般的に平均値 μ_i 、分散 σ_i^2 は空間位置 y_i の、相関係数 ρ_{ij} は地点 y_i と y_j の関数となる。

以上の事前確率場の仮定の下に観測を行い、 $n-1$ 地点で観測値 $x_i(i=1\sim n-1)$ が得られたとする。ここでは既知の観測値ベクトル $X_{n-1} = \{x_1, x_2, \dots, x_{n-1}\}^T$ を用いて未観測点における物理量 x_n の推定値を求め、その推定誤差あるいは観測更新程度について検討する。なお、条件付確率場とは事前情報に基づいてモデル化された確率場において、観測値が与えられた条件下で更新された事後の確率場のことを言う。

最適推定値については、事前確率分布を観測値を用いて更新した事後確率分布からエントロピーを定義し、それを最小にするようにして求める。推定誤差については無条件の事前確率場、観測値が得られた後の条件付事後確率場における確率密度関数から求められるエントロピーに基づいて間接的に評価する。

不規則に変動する物理量を確率的に配列した確率場において得られたサンプル実現値すなわち観測値は推定値であるので、確率変数 x と区別するため、 \underline{x} と下線を引いて示す。また推定式を表す変量は確率変数となるので推定量(estimator)、実際の観測値をこの推定式に

Table 1 Characteristics of unconditional stochastic fields having different distributions

Marginal distribution	Probability density function $f_X(x)$	Equivalent standard Gaussian transformation $g(z)$	Transformation coefficient F_{ij}
Gaussian	$\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$ ($-\infty < x < \infty$)	$\mu + \sigma z$	1
Lognormal	$\frac{1}{\sqrt{2\pi}\zeta x} \exp\left\{-\frac{1}{2}\left(\frac{\ln x - \lambda}{\zeta}\right)^2\right\}$ ($x > 0$)	$\exp(\lambda + \zeta z)$	$\frac{\ln(1 + \rho_{ij}\delta_i\delta_j)}{\rho_{ij}\sqrt{\ln(1 + \delta_i^2)\ln(1 + \delta_j^2)}}$
Shifted exponential	$\lambda \exp\{-\lambda(x-u)\}$ ($x \geq u$) 0 ($x < u$)	$u - \frac{1}{\lambda} \ln\{1 - \Phi(z)\}$	$1.229 - 0.367\rho_{ij} + 0.153\rho_{ij}^2$
Shifted Rayleigh	$\frac{x-u}{\alpha^2} \exp\left\{-\frac{1}{2}\left(\frac{x-u}{\alpha}\right)^2\right\}$ ($x \geq u$) 0 ($x < u$)	$u + \alpha[-2\ln\{1 - \Phi(z)\}]^{\frac{1}{2}}$	$1.028 - 0.029\rho_{ij}$
Gumbel (type I) Largest value	$\alpha \exp\{-\alpha(x-u)\} \exp\{-\exp\{-\alpha(x-u)\}\}$ ($-\infty < x < \infty$)	$u - \frac{1}{\alpha} \ln[-\ln\{\Phi(z)\}]$	$1.064 - 0.069\rho_{ij} + 0.005\rho_{ij}^2$
Uniform	$\frac{1}{b-a}$ ($a \leq x \leq b$) 0 ($x < a, x > b$)	$a + (b-a)\Phi(z)$	$1.047 - 0.047\rho_{ij}^2$

代入して得られる数値は確定値なので推定値 (estimate) と呼び、両者を明確に区別する。

3. 確率場における密度関数

本論文では正規分布のみならず非正規分布に従う条件付確率場の最大エントロピー推定法を提案する。エントロピーを定義するためには確率密度関数が必要である。そこで本章では5.の定式化に先立ち、無条件確率場の結合密度関数、条件付確率場の条件付密度関数の一般式を定義する。

(1) 等価標準正規確率変数への変換

非正規分布に従う確率場の結合確率密度関数は、与えられた非正規確率変数を等価な標準正規確率変数に変換し、変換場における標準正規結合確率密度関数を用いて求める。そこでこの等価標準正規確率変数への変換式を以下に示す。

確率変数 x_i の周辺確率分布関数 $F_{X_i}(x_i)$ は与えられた周辺確率密度関数 $f_{X_i}(x_i)$ より定義できる。この $F_{X_i}(x_i)$ は x_i に関して連続な増加関数である。確率変数ベクトル X_n は次式を満足するようにすると、等価標準正規変換によって標準正規確率変数 z_i よりなるベクトル $Z_n = \{z_1, z_2, \dots, z_n\}^T$ に変換できる。

$$F_{X_i}(x_i) = \Phi(z_i) \quad (i = 1 \sim n) \quad (1)$$

ここに、 $\Phi(\cdot)$ は標準正規確率分布関数である。

上式より、非正規確率変数 x_i は次式によって標準正規確率変数 z_i に関連づけられる。

$$\begin{aligned} x_i &= F_{X_i}^{-1}(\Phi(z_i)) \\ &= g(z_i) \end{aligned} \quad (2)$$

ただし、 $g(z)$ は z に関して非減少関数である。

Table 1 には6つの周辺分布 (正規分布, 対数正規分布, 指数分布, レイリー分布, グンベル分布, 一様分布) に関する $g(z)$ の具体例を示す。ここでは同表のように實際上重要と考えられる確率分布を対象にして確率場を仮定している。なお変換式には正規分布をも含めて定義する。

(2) 結合確率密度関数

X_n の結合確率密度関数 $f_{X_n}(X_n, R_n)$ は式(1)によると、

$$\begin{aligned} f_{X_n}(X_n, R_n) &= \frac{\partial(z_1, z_2, \dots, z_n)}{\partial(x_1, x_2, \dots, x_n)} \varphi_n(Z_n, R'_n) \\ &= \prod_{i=1}^n \frac{f_{X_i}(x_i)}{\varphi(z_i)} \varphi_n(Z_n, R'_n) \end{aligned} \quad (3)$$

と求められる。ただし、 $\varphi(\cdot)$ は標準正規確率密度関数、 $\varphi_n(Z_n, R'_n)$ は平均値ベクトル $\mathbf{0}$ 、分散ベクトル $\mathbf{1}$ 、相関係数行列 R'_n (後述) を有する n 次元の標準正規結合確率密度関数である。式(3)の分布形表示はNataf分布¹⁰⁾としてよく知られたものである。

式(1), (2)を用いると、式(3)の結合確率密度関数は関数 $g(z)$ の z に関する微分 $g'(z)$ を用いれば、次式で与えられる。

$$f_{X_n}(X_n, R_n) = \varphi_n(Z_n, R'_n) / \prod_{i=1}^n g'(z_i) \quad (4)$$

従って、非正規確率場における結合確率密度関数 $f_{X_n}(X_n, R_n)$ は等価標準正規変換式 $g(z_i)$ と相関係数行列 R'_n が与えられれば、簡単に求められることになる。この考え方は星谷・野田・稲田によって提案されたものである⁴⁾。

式(4)の評価に当っては R'_n が必要であるが、相関係数の定義式ならびに式(2), 式(4)の結合確率密度関数 $f_{X_2}(X_2, R_2)$ から, R_n の要素 ρ_{ij} は

$$\rho_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\frac{g(z_i) - \mu_i}{\sigma_i} \right) \left(\frac{g(z_j) - \mu_j}{\sigma_j} \right) \cdot \varphi_2(z_i, z_j, \rho'_{ij}) dz_j \quad (5)$$

と R'_n の要素 ρ'_{ij} と関連づけられる。

上式右辺は ρ'_{ij} の非線形式となっている。解析的に式(5)を解くことができる場合を除くと, ρ'_{ij} は一般的には収束計算によって求めなければならない。Liu and Kiureghian¹¹⁾はこの ρ'_{ij} を ρ_{ij} から算出する近似式を求めている。それによれば,

$$\rho'_{ij} = F_{ij} \rho_{ij} \quad (6)$$

として, ρ'_{ij} が得られる。 F_{ij} は変換係数であり, 相関係数 ρ_{ij} , 変動係数 $\delta_i (= \sigma_i / \mu_i)$ の関数である。例えば Table 1 の6タイプの分布形に関する変換係数を示すと, 同表右欄のようになる。正規分布, 対数正規分布については厳密解の, 他の分布についてはLiu and Kiureghianによる近似式の変換係数を示したものである。この近似式は $0.1 \leq \delta_i \leq 0.5$ の範囲で求められており, 指数分布における最大誤差は1.5%, 他の3つの分布ではほぼ厳密解を与える近似式となっている。

(3) 条件付確率密度関数

式(4)を用いると, $n-1$ 地点の観測点において確率変数 $x_i (i = 1 \sim n-1)$ が与えられた条件下における y_n 地点の確率変数 x_n の条件付確率密度関数は

$$f_{X_n}(x_n|c) = f_{X_n}(X_n, R_n) / f_{X_{n-1}}(X_{n-1}, R_{n-1}) \\ = \varphi(z_n|c) / g'(z_n) \quad (7)$$

と表わせる。ただし, 上記 $f_{X_n}(x_n|c)$ は $f_{X_n}(x_n|x_1, \dots, x_{n-1})$, $\varphi(z_n|c)$ は $\varphi(z_n|z_1, \dots, z_{n-1})$ を表わしており, それぞれ非正規確率変数, 標準正規確率変数の条件付確率密度関数であることを意味する。今後 $n-1$ 地点で確率変数 $\nu_i (i = 1 \sim n-1)$ が与えられたとき, 確率変数 ν_n の条件 $\nu_n|\nu_1, \dots, \nu_{n-1}$ を $\nu_n|c$ と示す。

星谷・野田・稲田⁴⁾は条件付非正規確率場の推定理論式を提案するのに, 式(7)を基本として用いている。本論でも条件付確率場のエントロピーを定義するのに, 式(7)の条件付確率密度関数を用いる。その考え方については5.で詳しく述べる。

$Z_{n-1} = \{z_1, z_2, \dots, z_{n-1}\}^T$ が与えられたときの z_n の条件付標準正規確率密度関数 $\varphi(z_n|c)$ は条件付平均値 $\mu_{z_n|c}$, 条件付分散 $\sigma_{z_n|c}^2$ を用いると, 次式のように表わすことができる。

$$\varphi(z_n|c) = \frac{1}{\sigma_{z_n|c}} \varphi\left(\frac{z_n - \mu_{z_n|c}}{\sigma_{z_n|c}}\right) \quad (8)$$

ただし, $\mu_{z_n|c}, \sigma_{z_n|c}^2$ はそれぞれ次のように求められる。

$$\mu_{z_n|c} = \sum_{i=1}^{n-1} \lambda'_i z_i \\ = \sum_{i=1}^{n-1} \lambda'_i g^{-1}(x_i) \quad (9)$$

$$\sigma_{z_n|c}^2 = 1 - \sum_{i=1}^{n-1} \lambda'_i \rho'_{in} \quad (10)$$

上式における重み係数 λ'_i は次式の連立方程式を解くことによって得られる。

$$\sum_{j=1}^{n-1} \lambda'_j \rho'_{ij} = \rho'_{in} \quad (i = 1 \sim n-1) \quad (11)$$

式(9)は標準正規確率場におけるKrigingの最適推定量(あるいは条件付平均値), 式(10)は同確率場における推定誤差分散(あるいは条件付分散)となっている。すなわち線形補間式は正規確率場においてのみ議論でき, このとき条件付平均値は最適推定量, 条件付分散は推定誤差分散と同一になる。

4. エントロピーと相互情報量

本章では5., 6.の理論式を提案する上で, また既往のエントロピーと本論のエントロピーの違いを明確にすることを目的に, 連続分布の確率変数のエントロピーと相互情報量に関して一般的な事柄^{5)~7)}を簡略的に示す。

(1) 情報量とエントロピー

観測点でサンプルを採取する際, 計測には何らかの誤差(ノイズ)を伴う。またモデル化に当たってもモデルの誤差を考慮しなければならない。その他自然界における現象にはノイズが必ず存在する。従ってノイズの大小に係わらず, 観測値はノイズに汚染されたと見なすのが適切である。

Shannonは, しばしば生起する事象の情報量は少なく, まれに生起する事象の情報量は多いという直感的感覚を基にして, 確率概念を使って情報量を1948年に定義した⁵⁾。この情報量とは情報を測る尺度のことを意味する。

情報量とエントロピーは同義語であるが, 明確に区別されるべきものである。すなわち, 情報量とはノイズを伴わない, エントロピーとはノイズを有する確率事象系に対して定義される⁷⁾。観測値はノイズを含むと考えるのが自然なので, ここでは情報量ではなくエントロピーなる用語を用いる。エントロピーとはもともと熱力学の分野で使われ始め, 不規則性を数量的に表現するために導入された。

(2) 離散値と連続値のエントロピー

今、有限事象系において事象 $a_i (i = 1 \sim n)$ の出現確率を $p(a_i)$ とすると、同事象のエントロピーは

$$-\log_2 p(a_i) \quad (12)$$

で定義でき、確率事象系 $A = (a_1, \dots, a_n)$ に対して定められる量

$$H_n(A) = E[-\log_2 p(a_i)] \\ = -\sum_{i=1}^n p(a_i) \log_2 p(a_i) \quad (13)$$

を平均エントロピーと区別して定義する。式(13)において $p(a_i)$ のどれかが0となる場合には $0 \log_2 0$ となる項が現われるが、この値は次の関係から0と約束する。

$$\lim_{x \rightarrow 0} x \log_2 x = 0 \quad (14)$$

式(13)は確率事象系 A の不確定性の度合を表わしている。 A が確定的な事象系(すなわち $p(a_i)$ のうち1つが1となる場合)では $H_n(A) = 0$ となり、 A が一様分布の場合(最も不確定な場合)には $H_n(A) = \log_2 n$ となって最大となる。

連続値を実数 x で表わし、その確率密度関数 $p(x)$ が定められているとする。実数全体 R の中で、分布 $X = (x, p(x); x \in R)$ に対し、

$$p_i = \int_{x_{i-1}}^{x_i} p(x) dx \quad (i = 1 \sim n) \quad (15)$$

とおくと、これは確率 p_i を有する有限事象系と考えることができる。このとき式(13)から平均エントロピーは

$$H_n(X) = -\sum_{i=1}^n p_i \log_2 p_i \quad (16)$$

となる。

$p(x)$ が連続関数で、 n が十分大きいときには $p_i = \Delta x p(x_i)$ と近似でき、平均エントロピーは次のようになる。

$$H_n(X) = -\sum_{i=1}^n \Delta x p(x_i) \log_2 (\Delta x p(x_i)) \\ = -\sum_{i=1}^n \{p(x_i) \log_2 p(x_i)\} \Delta x \\ - \sum_{i=1}^n p(x_i) \Delta x \log_2 \Delta x \quad (17)$$

分割数 n を増していくと $\Delta x \rightarrow 0$ となり、式(17)の右辺第2式第1項は

$$-\int_R p(x) \log_2 p(x) dx \quad (18)$$

に収束するが、第2項は

$$-\int_R p(x) dx \lim_{\Delta x \rightarrow 0} \log_2 \Delta x = -\lim_{\Delta x \rightarrow 0} \log_2 \Delta x \quad (19)$$

となり、無限大となる⁷⁾。

以上のことから連続分布の場合には式(13)と同様の方式で平均エントロピーを計算すると、その値は無限

大になってしまう。すなわち連続な量では取り得る値の自由度は無限にあるのだから、そこに含まれる情報量も無限であるはずなので、平均エントロピーは無限大になる。

式(19)の無限大項は意味がないのであろうか。連続量に対して平均エントロピーの概念は役に立たないのであろうか。実はそうではないのである。次節で示すように、平均相互情報量を定義すると、上述の無限大項が2つ、しかも一方は正の符号で、他方は負の符号で現われるので、これら無限大項は相互に相殺されるため、上述の発散は問題なくなるのである。

そこで連続分布の平均エントロピーは、離散的な場合の極限值としてではなく、形式的に次式で定義される⁵⁾。

$$H(X) = -\int_R p(x) \log_2 p(x) dx \quad (20)$$

上式は形式的に定義されたものであり、平均エントロピーとしての意味を十分にもっているとは言えない。従って $H(X)$ は負の値を示すこともあり得る。このことは有限事象系の場合と本質的に異なる。このようなことが生じるのは、式(20)の定義式において無限大に発散する項を除いてしまったことによる。

(3) 相互情報量

連続値をとる確率変数 x, y に対して周辺確率密度関数を $p(x), p(y)$ 、結合確率密度関数を $p(x, y)$ 、 Y が起こったときの X の条件付確率密度関数を $p(x|y)$ とする。

このとき $Y = \underline{y}$ の事象が与えられた条件下における X の条件付エントロピーは

$$\varepsilon(X|Y) = E[-\log_2 p(x|\underline{y})] \\ = -\int_R p(x|\underline{y}) \log_2 p(x|\underline{y}) dx \quad (21)$$

で、また事象系 Y に関する X の条件付平均エントロピーは

$$H(X|Y) = \int_R \varepsilon(X|Y) p(y) dy \\ = -\int_{R^2} p(x, y) \log_2 p(x|y) dx dy \quad (22)$$

で定義される。詳細については文献5), 7)を参照されたい。

もちろん式(21), (22)も $H(X)$ の形式的な定義と同様に条件付情報量としての意味を十分にもっているとは言えない。しかしながら、式(20), (22)で形式的に定義された $H(X), H(X|Y)$ が有界であれば、次式の $I(X; Y)$ は平均相互情報量としての意味を十分にもつ。

$$I(X; Y) = H(X) - H(X|Y) \geq 0 \quad (23)$$

無条件平均エントロピー $H(X)$ 、条件付平均エントロピー $H(X|Y)$ は必ずしも正の値とはならないことがあるが、平均相互情報量は負にならずに意味をもつ⁷⁾。す

なわち平均相互情報量は絶対的な尺度になり得るので、この意味で上述したように $I(X; Y)$ にはエントロピーなる用語を使わず、ここでは $I(X; Y)$ に対して情報量を冠した。

式(23)はShannonの不等式と呼ばれる⁷⁾。同式は、事象系 Y を知ることで、事象系 X の不確定さの量(情報量)が減少することはあっても増大することはないことを示している。このことは、事前情報の平均エントロピー $H(X)$ に対して、 Y を観測した後の X の事後情報の平均エントロピーを計量することにより、 X についての不確実性が減少することをエントロピーの概念を用いて定量的に計測できることを意味する。

平均相互情報量が大きいくほど、観測によって更新される程度は大きく、従って X の不確実性は大きいに減少することになる。すなわち平均相互情報量は推定誤差を評価する推定誤差分散に代わる指標として重要な意味をもっているのである。このことを条件付確率場の問題に当てはめて具体的に次章で詳述する。

5. 条件付確率場における最大エントロピー推定理論式の誘導

本章では、式(4)で示した無条件確率場の結合確率密度関数、式(7)の条件付確率密度関数を用い、無条件確率場において式(20)の無条件平均エントロピー、条件付確率場において式(21)の条件付エントロピー、式(22)の条件付平均エントロピーの定式化を示すとともに、最大エントロピー推定量の式を誘導する。またこれらの結果を不偏推定・最小誤差分散規範に基づく空間情報推定量の結果と比較分析することを試みる。

以下では条件付確率場として単一モードを有する分布形ならびに最も不確定な一様分布に注目した一般式の理論展開を行う。このような確率分布を取り扱うのは6.で示すような分布形が現実的に最も多く存在し、現象を説明するのに都合がよいためである。

(1) 最大エントロピー推定量

複数観測点で確率変数 x_i ($i = 1 \sim n-1$) または観測値 \hat{x}_i ($i = 1 \sim n-1$) が与えられた条件下で、未観測点における物理量 x_n の最大エントロピー推定量 \hat{x}_{E_n} または最大エントロピー推定値 \hat{x}_{E_n} を求める。ここで添字の E はエントロピーの概念に基づくことを意味する。

推定誤差を $a_n = x_n - \hat{x}_{E_n}$ とおき、その確率密度関数を $f_{A_n}(a_n)$ とする。その上で式(20)において $p(x)$ の代わりに $f_{A_n}(a_n)$ とおくと、平均エントロピーは $H(A_n) = -\int f_{A_n}(a_n) \log_2 f_{A_n}(a_n) da_n$ となる。この平均エントロピーを最小にする \hat{x}_{E_n} はKriging理論に基づく不偏最小誤差分散推定量に対応し、最小誤差エントロ

ピー推定量 \hat{x}_{E_n} になる。この定義に従えば、式(21)の条件付エントロピー、式(22)の条件付平均エントロピーは推定誤差を陽に表す指標となる。ここではこのような定義に基づくのではなく、文献(12), (13)と同様な定義に従って最大エントロピー推定量を求める。

連続分布のエントロピーは4.(2)の記号を用いると、式(12)から $-\log_2(\Delta x p(x_i))$ となる。これより、 $\Delta x \rightarrow 0$ の極限值を考えて無限大項を無視すると、連続量のエントロピーは $-\log_2 p(x)$ となる。そこで本論文では式(7)の条件付確率密度関数 $f_{X_n}(x_n|c)$ より、条件付確率場において次式のエントロピーを定義し、

$$-\log_2 f_{X_n}(x_n|c) \quad (24)$$

これを最小にする x_n を \hat{x}_{E_n} とおく。これは、観測点での変数が与えられた条件下において、事後の不確実性を最小にする x_n を見出すことを意味する。

文献(12), (13)では本論と同様にエントロピーに基づく推定法の概念を示している。それによれば、式(24)を最小にする、言い換えれば $f_{X_n}(x_n|c)$ を最大にする x_n を最大エントロピー推定量 \hat{x}_{E_n} と称している。エントロピーには4.で述べたように各種定義に従った述語が見られるので、本論文では既往文献(12), (13)の呼称・定義と対応させるため、 \hat{x}_{E_n} を最大エントロピー推定量と呼ぶことにする。

式(24)の最小解を得るためには

$$\left. \frac{1}{\ln 2} \frac{d \ln f_{X_n}(x_n|c)}{dx_n} \right|_{x_n = \hat{x}_{E_n}} = 0 \quad (25)$$

を満たす \hat{x}_{E_n} を求めればよい。上式は最小解の必要条件であるが、 \log 関数が上に凸であるため、式(24)の最小値を与える十分条件にもなっている。

式(25)は条件付確率密度関数の最大値すなわちモードを求めることに相当する。その解はMAP推定量(maximum a posteriori estimator)と呼ばれる。 $f_{X_n}(x_n)$ が一様分布であり、解 \hat{x}_{E_n} が一様分布の変数域の制限値すなわち境界値(Table 1のパラメータ a, b)をとらなければ、その解は最尤推定量と一致する。ただし後述の数値計算例でもわかるように、一様分布に従う条件付確率場において最大エントロピー推定値は境界値に至ることがある。従ってこの場合には最尤推定値にはならない。

式(7), (8)を用いると、式(25)より最大エントロピー推定量 \hat{x}_{E_n} は次式で表わされる。

$$\hat{x}_{E_n} = g(\hat{z}_{E_n}) \quad (26)$$

ただし、 \hat{z}_{E_n} は標準正規確率場での最適推定量であり、次式を満たす解である。

$$\hat{z}_{E_n} - \mu_{z_n|c} + \frac{g''(\hat{z}_{E_n})}{g'(\hat{z}_{E_n})} \sigma_{z_n|c}^2 = 0 \quad (27)$$

ここで、 $\mu_{z_n|c}, \sigma_{z_n|c}^2$ はそれぞれ式(9), (10)で求められる。

観測値 $\underline{x}_i (i = 1 \sim n-1)$ が与えられた条件下における最大エントロピー推定値 \hat{x}_{E_n} は式(29)を解いて得られる \hat{x}_{E_n} を式(28)に代入して求めればよい。

$$\hat{x}_{E_n} = g(\hat{z}_{E_n}) \quad (28)$$

$$\hat{z}_{E_n} - \mu_{z_n|c} + \frac{g''(\hat{z}_{E_n})}{g'(\hat{z}_{E_n})} \sigma_{z_n|c}^2 = 0 \quad (29)$$

ただし,
$$\mu_{z_n|c} = \sum_{i=1}^{n-1} \lambda_i g^{-1}(\underline{x}_i) \quad (30)$$

$$\sigma_{z_n|c}^2 = \sigma_{z_n|c}^2 \quad (31)$$

以上よりわかるように、本論で定義した最適推定量あるいは最適推定値の評価に当たっては不偏性の条件を用いていない。

(2) エントロピーに基づく更新程度の計量化

式(26)で求められた最大エントロピー推定量の2乗をとって $\varphi_{n-1}(\mathbf{Z}_{n-1}, \mathbf{R}'_{n-1})$ を掛け、 \mathbf{Z}_{n-1} で積分すると、 $E[\hat{x}_{E_n}^2]$ が得られる。従って \hat{x}_{E_n} に関連した推定誤差分散 $\sigma_{E_n}^2$ は $\mu_n^2 + \sigma_n^2 - E[\hat{x}_{E_n}^2]$ より算出することが可能である⁴⁾。しかしながらこの計算では後述の式(41)と同様に、解析的に積分可能な場合を除いて、一般には多重積分を行わねばならず、容易に解を求めることができない。そこで本論文では推定誤差分散を直接求めるのではなく、文献12)と同様に事前情報、観測を介して得られる情報量の変化を分析することにより、推定誤差を間接的に計量化することを試みる。

4.(3)の平均相互情報量において述べたように、式(26)の最大エントロピー推定量、式(28)の最大エントロピー推定値の推定精度は、事前情報に対し観測行為に伴って得られるエントロピーの減少を見ることによって評価できる。すなわち事前情報に基づく x_n のエントロピー $H(X_n)$ と、観測により更新された x_n の推定結果のエントロピー $H(X_n|C)$ を比較することにより、不確定性の減少度として推定誤差を間接的に計量できる。

まず無条件確率場において未観測点の物理量 x_n の平均エントロピーは式(20)より、

$$H(X_n) = - \int_R f_{X_n}(x_n) \log_2 f_{X_n}(x_n) dx_n \quad (32)$$

で表わせる。 R は確率変数 x_n の実対象領域を意味する。

観測値 \mathbf{X}_{n-1} が与えられたときの x_n の条件付エントロピーは式(21)の定義から、

$$\varepsilon(X_n|C) = - \int_R f_{X_n}(x_n|c) \log_2 f_{X_n}(x_n|c) dx_n \quad (33)$$

として求められる。ただし $x_n|c$ は $x_n|\underline{x}_1, \dots, \underline{x}_{n-1}$ の条件を意味する。文献12)では式(33)を最大エントロピー推定誤差と名付けている。

式(33)に式(7), (8)を代入すると、

$$\varepsilon(X_n|C) = \frac{1}{2 \ln 2} \left\{ \ln(2\pi e \sigma_{z_n|c}^2) \right.$$

$$\left. + 2 \int_{-\infty}^{\infty} \ln g'(\mu_{z_n|c} + \sigma_{z_n|c} y) \cdot \varphi(y) dy \right\} \quad (34)$$

となり、これより条件付確率場において観測事実が生じたとき、 x_n の条件付エントロピーが計算される。ただし $\mu_{z_n|c}, \sigma_{z_n|c}^2$ はおのこの式(30), (31)である。

一方式(33)を \mathbf{X}_{n-1} の結合確率分布で平均すると、 \mathbf{X}_{n-1} が起こったときの x_n に関する条件付平均エントロピーが求められる。式(22)より、

$$H(X_n|C) = \int_{R^{n-1}} \int \varepsilon(X_n|C) \cdot f_{\mathbf{X}_{n-1}}(\mathbf{X}_{n-1}, R_{n-1}) d\mathbf{X}_{n-1} \quad (35)$$

となるが、式(2)の等価標準正規変換式、式(4)の結合確率密度関数、式(7), (8)の条件付確率密度関数を式(34), (35)に用いると、式(35)は次式で表わせる。

$$H(X_n|C) = \frac{1}{2 \ln 2} \left\{ \ln 2(\pi e \sigma_{z_n|c}^2) \right. \\ \left. + 2 \int_R \ln g'(z_n) \cdot f_{X_n}(x_n) dx_n \right\} \quad (36)$$

推定誤差は厳密にはどの程度の推定精度を有するかを測る指標である。しかし本節の冒頭で述べたように最大エントロピー推定量の精度を数値的に求めることは容易でないので、ここでは観測に伴って生じるエントロピーの減少度を調べることによってこれを計量化する。この減少度は基本的には推定と無関係であるが、6.で具体的に述べるように、最大エントロピー推定量、条件付エントロピー、条件付平均エントロピーはKriging理論の最適推定量、条件付分散、推定誤差分散と関連している。そのためこの関連性を示すことは重要である。本節ではエントロピーに関連した一般式を示すに留めたが、この関係を示す具体例については6.で述べる。

4.(3)で述べたように、Shannonの不等式から事後確率分布の条件付平均エントロピー $H(X_n|C)$ は事前確率分布に関する無条件平均エントロピー $H(X_n)$ よりも大きくならないことが理論的に保証される。すなわちデータが増えるに従い、事前に設定した不確かさの量は減少していくことを意味する。ただし観測結果が得られた後に式(34)で示す $\varepsilon(X_n|C)$ が $H(X_n)$ より大きくなることはあり得る。

ところで、未観測点すなわち推定点 y_n が観測点の一つ $y_i (i = 1 \sim n-1)$ と一致すると、重み係数 λ_i の性質から式(9)の $\mu_{z_n|c}$ 、式(30)の $\mu_{z_n|c}$ は z_i, \underline{z}_i に、式(10)の $\sigma_{z_n|c}^2$ 、式(31)の $\sigma_{z_n|c}^2$ は0に至るので、条件付確率密度関数 $f_{X_n}(x_n|c), f_{X_n}(x_n|\underline{x}_i)$ はそれぞれ $\delta(x_n - x_i), \delta(x_n - \underline{x}_i)$ で表わされる。 x_n が x_i, \underline{x}_i に一致すると、これらは ∞ の値となる。このとき式(33)の被積分関数は $\delta(x_n - \underline{x}_i) \log_2 \delta(x_n - \underline{x}_i)$ となるので、同式の $\varepsilon(X_n|C)$ は $-\infty$ となる。同様に式(35)の $H(X_n|C)$ も $-\infty$ の値になる。

さらに、 \hat{x}_{E_n} は式(26)の定義より x_i を与え、式(28)より推定値 \hat{x}_{E_n} は \underline{x}_i となる。このように推定量、推定値が観測点の確率変数、観測値と完全に一致することにより、条件付エントロピー $-\varepsilon(X_n|C)$ 、条件付平均エントロピー $H(X_n|C)$ は $-\infty$ で表わせることになる。

一方、観測点から遠く離れた未観測点では観測値の影響が小さくなり、式(11)の ρ'_{in} ($i = 1 \sim n-1$) は0に近づくことになる。このとき、観測点間の相関係数 ρ'_{ij} ($i, j = 1 \sim n-1$) の有限値に対し、式(11)の λ'_i ($i = 1 \sim n-1$) は0に近づく。従って $\mu_{z_n|c}$ 、 $\mu_{z_n|c}$ は0、 $\sigma_{z_n|c}^2$ 、 $\sigma_{z_n|c}^2$ は1に至る。このことから式(7)の $f_{X_n}(x_n|c)$ は $f_{X_n}(x_n)$ に漸近するので、式(26)の解 \hat{x}_{E_n} は結局無条件確率密度関数の最大値を与えるモードとなる。このモードは正規分布では無条件平均値と一致するが、一様分布以外の非正規分布では無条件平均値よりも小さい。その具体的な諸量については6.で詳述する。なお一様分布のモードは特定できない。

(3) 星谷・野田・稲田の推定理論

本論で提案したエントロピーに基づく最適推定理論式を不偏推定・最小誤差分散規範によるKriging理論の結果と比較するため、以下には星谷・野田・稲田の推定理論式⁴⁾を示す。

まず条件付非正規確率場における未観測点 y_n の物理量 x_n の条件付平均値 $\mu_{n|c}$ 、条件付分散 $\sigma_{n|c}^2$ は式(7)、(8)の条件付確率密度関数を用いると、それぞれ次のようになる。

$$\mu_{n|c} = \int_{-\infty}^{\infty} g(z_n)\varphi(z_n|c)dz_n \quad (37)$$

$$\sigma_{n|c}^2 = \int_{-\infty}^{\infty} g^2(z_n)\varphi(z_n|c)dz_n - \mu_{n|c}^2 \quad (38)$$

ここに、式(37)は不偏推定式であり、同時に最小誤差分散式を満足する。従って式(37)は最適推定量 \hat{x}_{K_n} となる。添字の K はKrigingの推定規範に基づくことを示している。

なお、観測値 X_{n-1} が与えられた条件下で未観測点 y_n における最適推定値 \hat{x}_{K_n} 、条件付分散 $\sigma_{n|c}^2$ はそれぞれ次の式で表わせる。

$$\begin{aligned} \hat{x}_{K_n} &= \mu_{n|c} \\ &= \int_{-\infty}^{\infty} g(z_n)\varphi(z_n|c)dz_n \end{aligned} \quad (39)$$

$$\sigma_{n|c}^2 = \int_{-\infty}^{\infty} g^2(z_n)\varphi(z_n|c)dz_n - \mu_{n|c}^2 \quad (40)$$

ここで、式(39)、(40)の $\varphi(z_n|c)$ は、式(8)において、 $\mu_{z_n|c}$ の代わりに式(30)の $\mu_{z_n|c}$ を、 $\sigma_{z_n|c}^2$ の代わりに式(31)の $\sigma_{z_n|c}^2$ を用いて計算すればよい。

一方、推定誤差分散 $\sigma_{K_n}^2$ は定義より、

$$\sigma_{K_n}^2 = E[(x_n - \hat{x}_{K_n})^2]$$

$$\begin{aligned} &= \mu_n^2 + \sigma_n^2 \\ &\quad - \int_{R^{n-1}} \int \left(\int_{-\infty}^{\infty} g(z_n)\varphi(z_n|c)dz_n \right)^2 \\ &\quad \cdot \varphi_{n-1}(Z_{n-1}, R'_{n-1})dZ_{n-1} \end{aligned} \quad (41)$$

で求められる。

以上のことから、1)最適推定値は条件付平均値に一致すること、2)非正規確率場における推定誤差分散は条件付分散とは異なること、3)最適推定値と条件付分散は式(39)、(40)において観測値によって決められる $\mu_{z_n|c}$ を評価式に含むため、観測値に依存すること、4)推定誤差分散は観測点配置にのみ依存し、観測値とは独立であること、5)推定誤差分散は無条件分散 σ_n^2 よりも大きくはならないこと、6)非正規確率場において式(40)の条件付分散は無条件分散より大きくなることもあり得ることなどが星谷ら⁴⁾によって明らかにされた。なお正規確率場において、推定誤差分散と条件付分散は同一値となることに注意する必要がある。

既往論文の中には上記事実を正しく理解せず、対数正規確率場の推定問題に対し、最適推定量と最適推定値を混同し、推定誤差分散を条件付分散と見なしたり、Krigingの定義に従って最適推定値と推定誤差分散を正確に評価していなかったりする場合がある。このような不相应な解釈はこれまで条件付正規確率場のKriging理論が主に多用されてきたため、条件付正規確率場と条件付非正規確率場の本質的な相違点を論理的にきちんと把握していないことを原因としている。

これまでの理論展開式より、エントロピー規範に基づく式(26)～(31)、(34)、(36)はそれぞれ不偏推定・最小誤差分散規範に基づく式(37)、(39)、(40)、(41)に対応していることがわかる。これより、非正規確率場において星谷らによって指摘された特徴はエントロピー理論においても成立していることが理論的に明らかになった。従って上述のように現象の解明に当たって明らかに不相应な理論を当てはめて構成された既往の研究結果は正しくないことが最大エントロピー推定法の観点からも証明されたと言える。

6. 提案手法の任意確率場への適用

前章で提案した手法をTable 1に示す6つのタイプの分布形(正規分布、対数正規分布、指数分布、レイリー分布、グンベル分布、一様分布)に適用し、具体的に最大エントロピー推定値(式(28))、無条件平均エントロピー(式(32))、条件付エントロピー(式(34))、条件付平均エントロピー(式(36))を誘導した結果を以下に示す。なお各分布は同一の平均値 μ 、分散 σ^2 を有すると仮定しているため、各パラメータはTable 2の関係のように

Table 2 Parameters of marginal distribution function

Marginal distribution	Parameters (Mean = μ , Variance = σ^2)
Gaussian	μ σ
Lognormal	$\zeta = \sqrt{\ln(1 + \delta^2)}$ $\lambda = \ln \mu - \frac{\zeta^2}{2}$
Shifted exponential	$\lambda = \frac{1}{\sigma}$ $u = \mu - \sigma$
Shifted Rayleigh	$\alpha = \sigma / \sqrt{2 - \frac{\pi}{2}}$ $u = \mu - \sqrt{\frac{\pi}{2}} \alpha$
Gumbel (type I) Largest value	$\alpha = \frac{\pi}{\sqrt{6}\sigma}$ $u = \mu - \frac{\sqrt{6}\gamma}{\pi} \sigma$
Uniform	$a = \mu - \sqrt{3}\sigma$ $b = \mu + \sqrt{3}\sigma$

なる。

(1) 正規分布

まず点 y_i における観測値 $x_i (i = 1 \sim n-1)$ が与えられた条件下で未観測点 y_n における最大エントロピー推定値は式 (28) ~ (31) から次式で表わせる。

$$\hat{x}_{E_n} = \mu_n + \sum_{i=1}^{n-1} a_i (x_i - \mu_i) \quad (42)$$

ただし、重み係数 a_i は、共分散 C_{ij} を用いると、次式の連立方程式により与えられる。

$$\sum_{j=1}^{n-1} a_j C_{ij} = C_{in} \quad (i = 1 \sim n-1) \quad (43)$$

無条件平均エントロピー $H(X_n)$ は式 (32) より、

$$H(X_n) = \ln(2\pi e \sigma_n^2) / 2 \ln 2 \quad (44)$$

条件付エントロピー $\varepsilon(X_n|C)$ は式 (34) より、

$$\varepsilon(X_n|C) = \ln(2\pi e \sigma_{x_n|c}^2) / 2 \ln 2 \quad (45)$$

条件付平均エントロピー $H(X_n|C)$ は式 (36) より、

$$H(X_n|C) = \ln(2\pi e \sigma_{x_n|c}^2) / 2 \ln 2 \quad (46)$$

と求められる。ただし、次式が成立する。

$$\begin{aligned} \sigma_{x_n|c}^2 &= \sigma_{x_n|c}^2 \\ &= \sigma_n^2 - \sum_{i=1}^{n-1} a_i C_{in} \end{aligned} \quad (47)$$

式 (42), (47) はそれぞれ条件付正規確率場における条件付平均値 (あるいは最適推定値), 条件付分散 (あるいは推定誤差分散) と同一の式を与える。式 (45), (46) は式中に観測値を含まないために一致する。条件付平均エントロピー $H(X_n|C)$ は推定誤差分散 $\sigma_{x_n|c}^2$ の対数値

をとっただけなので、基本的には $\sigma_{x_n|c}^2$ の性質と同等の振舞をする。よって条件付正規確率場における最大エントロピー推定結果は Kriging の推定結果と等価であることが理解できる。

Kriging の推定値は条件付平均値である。条件付確率密度関数 $f_{X_n}(x_n|c)$ は正規分布に従うので、その条件付平均値は x_n に関して $\log_2 f_{X_n}(x_n|c)$ の関数を最大化する値となる。このため、最大エントロピー推定値は Kriging の最適推定値と同一なる。

$\sigma_{x_n|c}^2$ は無条件分散 σ_n^2 に等しいかそれよりも小さいため、式 (44) と式 (46) の比較からわかるように、 $H(X_n|C)$ は $H(X_n)$ よりも大きくはならない。観測点よりもかなり遠方の未観測点において推定誤差分散を求めると、観測値の影響が小さくなるので、式 (47) の推定誤差分散は無条件分散に近づく。このようなとき、 $H(X_n|C)$ は $H(X_n)$ に等しくなる。

未観測点 y_n が観測点の一つ y_i に一致すると、式 (43) の重み係数 a_i の性質から、式 (42) の最大エントロピー推定値は観測値 x_i に一致し、式 (47) の推定誤差分散は 0 になる。このとき式 (46) の条件付平均エントロピーは 5. (2) で示したように、 $f_{X_n}(x_n|c)$ が $\delta(\cdot)$ 関数で表わせるので、 $-\infty$ の値をとる。

平均相互情報量は、式 (23) よりわかるように式 (44) から式 (46) を引いて次式のように求められ、Shannon の不等式が成立していることがわかる。

$$I(X_n; C) = H(X_n) - H(X_n|C) \geq 0 \quad (48)$$

式 (48) の定義を用いると、条件付確率場における任意地点で有界な $I(X_n; C)$ を評価することにより、その値の大小から観測に伴う事前確率場の更新程度、すなわち計測で得られる情報の相対的な大きさを測ることが可能になる。この重要な性質は以下の非正規分布においても成立する。

(2) 対数正規分布

最大エントロピー推定値は式 (28) ~ (31) より、次式で表わせる。

$$\hat{x}_{E_n} = \exp(\mu_{K_{en}} - \sigma_{K_{en}}^2) \quad (49)$$

ただし、 $\mu_{K_{en}}, \sigma_{K_{en}}^2$ はそれぞれ $\ln x$ の正規確率場における条件付平均値 (最適推定値), 条件付分散 (推定誤差分散) であり、次式で求められる。

$$\mu_{K_{en}} = \lambda_n + \sum_{i=1}^{n-1} b_i (\ln x_i - \lambda_i) \quad (50)$$

$$\sigma_{K_{en}}^2 = \zeta_n^2 - \sum_{i=1}^{n-1} b_i C'_{in} \quad (51)$$

なお、 λ_n, λ_i は $\ln x_n, \ln x_i$ の平均値、 ζ_n^2 は $\ln x_n$ の分散である。また、重み係数 b_i は $\ln x_i$ と $\ln x_j$ の共分散 C'_{ij} を

用いると、次式の連立方程式の解として与えられる。

$$\sum_{j=1}^{n-1} b_j C'_{ij} = C'_{in} \quad (i = 1 \sim n-1) \quad (52)$$

無条件平均エントロピー $H(X_n)$, 条件付エントロピー $\varepsilon(X_n|\underline{C})$, 条件付平均エントロピー $H(X_n|C)$ はそれぞれ式(32), (34), (36)より次のようになる。

$$H(X_n) = \frac{1}{2 \ln 2} \{2\lambda_n + \ln(2\pi e \zeta_n^2)\} \\ = \ln(2\pi e \mu_n^2 \zeta_n^2 e^{-\zeta_n^2}) / 2 \ln 2 \quad (53)$$

$$\varepsilon(X_n|\underline{C}) = \frac{1}{2 \ln 2} \{2\mu_{K_{en}} + \ln(2\pi e \sigma_{K_{en}}^2)\} \\ = \ln(2\pi e \hat{\zeta}_{E_n}^2 \sigma_{K_{en}}^2 e^{2\sigma_{K_{en}}^2}) / 2 \ln 2 \quad (54)$$

$$H(X_n|C) = \frac{1}{2 \ln 2} \{2\lambda_n + \ln(2\pi e \sigma_{K_{en}}^2)\} \\ = \ln(2\pi e \mu_n^2 \sigma_{K_{en}}^2 e^{-\zeta_n^2}) / 2 \ln 2 \quad (55)$$

これより、1) $\hat{\zeta}_{E_n}$, $\varepsilon(X_n|\underline{C})$ が観測値に依存すること、2) $H(X_n|C)$ は観測値に無関係であること、3) $I(X_n; C)$ は $\ln 2 \{ \pi e (\zeta_n^2 - \sigma_{K_{en}}^2) \} / 2 \ln 2$ で与えられるので、観測を通して平均相互情報量が増加すること、4) 観測点と未観測点が一致すると、 $\mu_{K_{en}}$ は $\ln \underline{x}_i$ に、 $\sigma_{K_{en}}^2$ は 0 になるので、 $\hat{\zeta}_{E_n}$ は観測値 \underline{x}_i と一致し、先の正規分布と同様に $\varepsilon(X_n|\underline{C})$, $H(X_n|C)$ は式(33), (35)の定義式から $-\infty$ に至ることがわかる。

観測点から離れた遠方地点において、 $\mu_{K_{en}}$ は λ_n に、 $\sigma_{K_{en}}^2$ は ζ_n^2 に近づくため、最適推定値 $\hat{\zeta}_{E_n}$ は $\hat{\zeta}_{E_\infty} = \exp(\lambda_n - \zeta_n^2)$ となる。これは無条件平均値 $\exp(\lambda_n + \zeta_n^2/2)$ とは異なり、小さ目の値である。またこのとき遠方場の $\varepsilon(X_n|\underline{C})$, $H(X_n|C)$ は $H(X_n)$ に近づく。

一方、条件付対数正規確率場における最適推定値 $\hat{\zeta}_{K_n}$, 条件付分散 $\sigma_{n|\underline{c}}^2$, 推定誤差分散 $\sigma_{K_n}^2$ はおのおの次式で与えられる⁴⁾。

$$\hat{\zeta}_{K_n} = \exp\left(\mu_{K_{en}} + \frac{1}{2}\sigma_{K_{en}}^2\right) \quad (56)$$

$$\sigma_{n|\underline{c}}^2 = \hat{\zeta}_{K_n}^2 \left(e^{\sigma_{K_{en}}^2} - 1\right) \quad (57)$$

$$\sigma_{K_n}^2 = \mu_n^2 e^{\zeta_n^2} \left(1 - e^{-\sigma_{K_{en}}^2}\right) \quad (58)$$

これより、条件付正規確率場の性質と異なり、1) 条件付分散と推定誤差分散は同一値を示さないこと、2) 最適推定値と条件付分散は観測値に依存するが、推定誤差分散は観測値に独立であることがわかる。

式(49)と式(56), 式(54)と式(57), 式(55)と式(58)を比較するとわかるように、最大エントロピー推定法による結果は不偏最小誤差分散規範に基づく推定法の結果と微妙に異なっている。6.(1)で述べた条件付正規確率場の如く、両手法による結果は完全な等価性を示さないが、お互いに関連していることが理解できる。

式(49)の最大エントロピー推定値は式(56)に示す Kriging の最適推定値のように観測点を通して滑らかに

推定されるのではなく、小さ目の値となる。これは遠方場の推定値 $\hat{\zeta}_{E_\infty}$ すなわち無条件確率場のモードに近くように最大エントロピー推定値が評価されるため、不偏性条件を考慮していないことによる。このモード値は無条件平均値よりも小さい。

(3) 指数分布

最大エントロピー推定値 $\hat{\zeta}_{E_n}$ は、式(29)を解いて求められる $\hat{\zeta}_{E_n}$ (すなわち式(60))を式(28)(すなわち式(59))に代入して得られる。

$$\hat{\zeta}_{E_n} = u_n - \frac{1}{\lambda_n} \ln \{1 - \Phi(\hat{\zeta}_{E_n})\} \quad (59)$$

$$\text{ここに、} \hat{\zeta}_{E_n} = \frac{\mu_{z_n|\underline{c}} - \lambda_n \sigma_{z_n|\underline{c}}^2}{1 - \sigma_{z_n|\underline{c}}^2} \quad (60)$$

$$\mu_{z_n|\underline{c}} = \sum_{i=1}^{n-1} \lambda_i \Phi^{-1} \left[1 - \exp\{\lambda_i(u_i - \underline{x}_i)\} \right] \quad (61)$$

無条件平均エントロピー $H(X_n)$ は

$$H(X_n) = \ln(e^2 \sigma_n^2) / 2 \ln 2 \quad (62)$$

となり、条件付エントロピー $\varepsilon(X_n|\underline{C})$ は

$$\varepsilon(X_n|\underline{C}) = \frac{1}{2 \ln 2} \left[\ln(e^2 \sigma_n^2 \sigma_{z_n|\underline{c}}^2) - \mu_{z_n|\underline{c}}^2 - \sigma_{z_n|\underline{c}}^2 \right. \\ \left. - 1 - 2 \int_{-\infty}^{\infty} \ln\{1 - \Phi(\mu_{z_n|\underline{c}} + \sigma_{z_n|\underline{c}} y)\} \right. \\ \left. \cdot \varphi(y) dy \right] \quad (63)$$

で、条件付平均エントロピーは

$$H(X_n|C) = \ln(e^2 \sigma_n^2 \sigma_{z_n|\underline{c}}^2) / 2 \ln 2 \quad (64)$$

で与えられる。

式(63)の評価式には1次元全無限区間積分を含む。そこで後述の数値計算例においては、この積分項を計算するため、高橋・森の2重指数関数型積分公式に基づく自動積分法¹⁴⁾を採用する。

観測点と推定点が一致すると、定義から最大エントロピー推定値は観測値に、条件付エントロピー、条件付平均エントロピーは $-\infty$ になる。遠方場では $\mu_{z_n|\underline{c}} \rightarrow 0$, $\sigma_{z_n|\underline{c}}^2 \rightarrow 1$ となるため、最大エントロピー推定値は無条件平均値 $u_n + 1/\lambda_n$ より小さい u_n に近づき、内挿点の推定値はこの u_n に近くように求められる。また遠方場において式(63), (64)は $H(X_n)$ に近づくことが証明される。

(4) レイリー分布

最大エントロピー推定値 $\hat{\zeta}_{E_n}$ は式(28)~(31)より定式化すると、次の実超越方程式の解として与えられる。

$$\Phi^{-1} \left[1 - \exp \left\{ -\frac{1}{2} \left(\frac{\hat{\zeta}_{E_n} - u_n}{\alpha_n} \right)^2 \right\} \right] - \frac{1}{1 - \sigma_{z_n|\underline{c}}^2}$$

$$\left\{ \mu_{z_n|E} - \left(\frac{\hat{x}_{E_n} - u_n}{\alpha_n^2} - \frac{1}{\hat{x}_{E_n} - u_n} \right) \sigma_{z_n|E}^2 \right\} = 0 \quad (65)$$

ここに,

$$\mu_{z_n|E} = \sum_{i=1}^{n-1} \lambda_i \Phi^{-1} \left[1 - \exp \left\{ -\frac{1}{2} \left(\frac{x_i - u_i}{\alpha_i} \right)^2 \right\} \right] \quad (66)$$

式(65)を解くのに、ここではMuller法¹⁴⁾を用いる。この方法は求めようとする根の3つの近似値を用いて、実超越方程式 $f(x)$ を2次のニュートン補間多項式 $P(x)$ で近似し、 $P(x) = 0$ の根の一方を $f(x)$ の次の近似根として採用することにより、次々と反復的に近似していくものである。

無条件平均エントロピーを式(32)に従って算出すると、

$$\frac{1}{2 \ln 2} \left[\ln \left\{ e^2 \sigma_n^2 / \left(2 - \frac{\pi}{2} \right) \right\} - 2 \int_0^\infty y \ln y \cdot e^{-\frac{y^2}{2}} dy \right] \quad (67)$$

と求められる。上式の積分項は有限値として定義できないので、連続値のエントロピーにおいて形式的に無限大項を除くと同様の意味で、ここでは次式によって無条件平均エントロピーを形式的に定義する。なお条件付平均エントロピーにも同様の半無限区間積分が現われ相殺されるので、このように積分項を割愛してエントロピーを定義しても、平均相互情報量の評価に当たっては十分に意味をもつことになる。

$$H(X_n) = \ln \left\{ e^2 \sigma_n^2 / \left(2 - \frac{\pi}{2} \right) \right\} / 2 \ln 2 \quad (68)$$

式(34)によって条件付エントロピー $\varepsilon(X_n|C)$ を記すと、次のようになる。

$$\begin{aligned} \varepsilon(X_n|C) = & \frac{1}{2 \ln 2} \left[\ln \left\{ e^2 \sigma_n^2 \sigma_{z_n|E}^2 / \left(2 - \frac{\pi}{2} \right) \right\} \right. \\ & - \mu_{z_n|E}^2 - \sigma_{z_n|E}^2 - 1 \\ & - \int_{-\infty}^{\infty} \ln \left[-2 \ln \{ 1 - \Phi(\mu_{z_n|E} + \sigma_{z_n|E} y) \} \right] \cdot \varphi(y) dy \\ & \left. - 2 \int_{-\infty}^{\infty} \ln \{ 1 - \Phi(\mu_{z_n|E} + \sigma_{z_n|E} y) \} \cdot \varphi(y) dy \right] \quad (69) \end{aligned}$$

上式の2つの積分項のうち最初に現れる項は式(67)の積分項に対応する項である。従って実際の計算においてはこの項を除いて $\varepsilon(X_n|C)$ を評価する。このときそれを X_{n-1} の結合確率密度関数を用いて平均すると、 X_{n-1} が起こったときの x_n の条件付平均エントロピーは

$$H(X_n|C) = \ln \left\{ e^2 \sigma_n^2 \sigma_{z_n|E}^2 / \left(2 - \frac{\pi}{2} \right) \right\} / 2 \ln 2 \quad (70)$$

となる。

観測点から離れた遠方場において、式(65)から最大エントロピー推定値は $u_n + \alpha_n$ になる。この値は無条件平均値 $u_n + \sqrt{\pi/2} \alpha_n$ よりも小さい。このとき条件付エントロピー、条件付平均エントロピーは無条件平均エントロピーに近づくことが示される。

(5) グンベル分布

最大エントロピー推定値 \hat{x}_{E_n} は次の実超越方程式の根として与えられる。

$$\begin{aligned} \Phi^{-1} \left\{ e^{-e^{-\alpha_n (\hat{x}_{E_n} - u_n)}} \right\} - \frac{1}{1 - \sigma_{z_n|E}^2} \left[\mu_{z_n|E} \right. \\ \left. - \alpha_n \left\{ 1 - e^{-\alpha_n (\hat{x}_{E_n} - u_n)} \right\} \sigma_{z_n|E}^2 \right] = 0 \quad (71) \end{aligned}$$

$$\text{ただし、} \mu_{z_n|E} = \sum_{i=1}^{n-1} \lambda_i \Phi^{-1} \left\{ e^{-e^{-\alpha_i (x_i - u_i)}} \right\} \quad (72)$$

無条件平均エントロピーを求めると、次のようになる。

$$H(X_n) = \left\{ \ln \left(\frac{6e^2 \sigma_n^2}{\pi^2} \right) + 2\gamma \right\} / 2 \ln 2 \quad (73)$$

ここに、 γ はEuler定数である。

条件式エントロピー $\varepsilon(X_n|C)$ は式(34)より、次のように表わせる。

$$\begin{aligned} \varepsilon(X_n|C) = & \frac{1}{2 \ln 2} \left[\ln \left(\frac{6e^2 \sigma_n^2 \sigma_{z_n|E}^2}{\pi^2} \right) - \mu_{z_n|E}^2 - \sigma_{z_n|E}^2 \right. \\ & - 1 - 2 \int_{-\infty}^{\infty} \ln \Phi(\mu_{z_n|E} + \sigma_{z_n|E} y) \cdot \varphi(y) dy \\ & \left. - 2 \int_{-\infty}^{\infty} \ln \{ -\ln \Phi(\mu_{z_n|E} + \sigma_{z_n|E} y) \} \cdot \varphi(y) dy \right] \quad (74) \end{aligned}$$

また条件付平均エントロピー $H(X_n|C)$ を求めると、次のように表わせる。

$$H(X_n|C) = \left\{ \ln \left(\frac{6e^2 \sigma_n^2 \sigma_{z_n|E}^2}{\pi^2} \right) + 2\gamma \right\} / 2 \ln 2 \quad (75)$$

推定点と観測点が一致すると、最大エントロピー推定値 \hat{x}_{E_n} は観測値に完全に一致する。その結果条件付エントロピー $\varepsilon(X_n|C)$ 、条件付平均エントロピー $H(X_n|C)$ は $-\infty$ になることが証明される。一方、観測点から離れた遠方場に位置する推定点では \hat{x}_{E_n} が u_n (無条件平均値の $u_n + \gamma/\alpha_n$ と異なる)に、式(74)の $\varepsilon(X_n|C)$ 、式(75)の $H(X_n|C)$ が $H(X_n)$ に近づくことも導出される。

(6) 一様分布

最大エントロピー推定値は次式で表わせる。

$$\hat{x}_{E_n} = (b_n - a_n) \Phi \left(\frac{\mu_{z_n|E}}{1 - \sigma_{z_n|E}^2} \right) + a_n \quad (76)$$

$$\text{ただし, } \mu_{z_n|E} = \sum_{i=1}^{n-1} \lambda_i \Phi^{-1} \left(\frac{x_i - a_i}{b_i - a_i} \right) \quad (77)$$

無条件確率場の平均エントロピーは

$$H(X_n) = \ln(12\sigma_n^2)/2 \ln 2 \quad (78)$$

となり, 条件付確率場におけるエントロピー $\varepsilon(X_n|C)$ は

$$\varepsilon(X_n|C) = \frac{1}{2 \ln 2} \left\{ \ln(12\sigma_n^2 \sigma_{z_n|E}^2) - \mu_{z_n|E}^2 - \sigma_{z_n|E}^2 + 1 \right\} \quad (79)$$

で, 条件付平均エントロピーは

$$H(X_n|C) = \ln(12\sigma_n^2 \sigma_{z_n|C}^2)/2 \ln 2 \quad (80)$$

で求められる。

一様分布のモードを与える変数値は境界値 a_n と b_n の間で任意である。このため観測値の影響を受け, 推定値は一様分布の境界値をとることがある。従って, 観測点から離れた内挿点では $\mu_{z_n|E} \rightarrow 0$, $\sigma_{z_n|E}^2 \rightarrow 1$ となるため, 最大エントロピー推定値は2つの境界値間で任意値をとりえる。そこでここでは観測点から離れた遠方場において, \hat{x}_{E_n} は無条件平均値 $(a_n + b_n)/2$ に近づくものと定義する。なお, 遠方場の $\varepsilon(X_n|C)$, $H(X_n|C)$ は $H(X_n)$ になる。

推定点と観測点が一致すると, 定義式ならびに極限計算から, 最大エントロピー推定値は観測値と完全に一致し, 条件付エントロピー, 条件付平均エントロピーは $-\infty$ になることが理解できる。

7. 数値計算例および考察

(1) 前提条件

Kriging法との対比から提案手法の性質を明らかにする上で, ここでは簡単のために1次元均一確率場を対象とする。対象とした確率分布はTable 1, Table 2に示すように, 正規分布と5つの非正規分布(対数正規分布, 指数分布, レイリー分布, グンベル分布, 一様分布)である。6つの周辺分布の無条件平均値は $\mu=2.5$, 無条件分散は $\sigma^2=1.5625$ (無条件変動係数 $\delta=0.5$)として, 同一値を与える。このとき各周辺分布のパラメータはTable 2のようになる。また x_i と x_j の無条件相関係数 ρ_{ij} は次式に従うと仮定する。

$$\rho_{ij} = \exp(-\|y_i - y_j\|/4) \quad (81)$$

ただし, y_i, y_j は i 点, j 点の座標である。

本論では割愛したが, 事前分布として与えた各周辺分布の無条件確率密度関数は文献4)の図-7と同一である。観測点は8点($y_i=0, 10, 20, 40, 50, 60, 90, 100$)とした。観測点における観測値は, 相関を有する無条件正規確率場の同時シミュレーションによって予め求め

たもので, 各周辺分布について同一値を設定した。分布形によってはここで発生した観測値がまれにしか生じないケースもあるが, この設定条件は本論の分析結果に影響を与えるものではない。

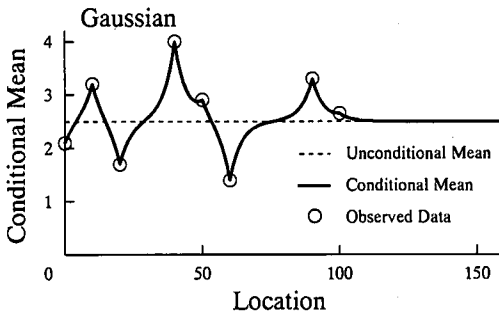
(2) Kriging法の推定結果

最大エントロピー推定法による推定結果を議論する前に, Kriging法の結果を示す。文献4)では条件付シミュレーションを行い, 分布形の相違が条件付確率場の推定結果に及ぼす影響を検討した。しかしながら同論文では理論解を求めていなかったため, ここでは式(39)の条件付平均値(あるいは最適推定値), 式(40)の条件付分散の理論解をFig.1, Fig.2に, 5,000回のサンプル場を発生して得られた結果からサンプル統計量(条件付サンプル平均値, 条件付サンプル分散値)を求めてFig.3, Fig.4に示す。なお, 条件付正規確率場, 条件付対数正規確率場の理論解は式(42), (47), (56), (57)のように求められる。しかしながらここでは, 他の分布形と同様に, 式(39), (40)の全無限区間積分を高橋・森の2重指数関数型積分公式に基づく自動積分法¹⁴⁾を用いて計算した。

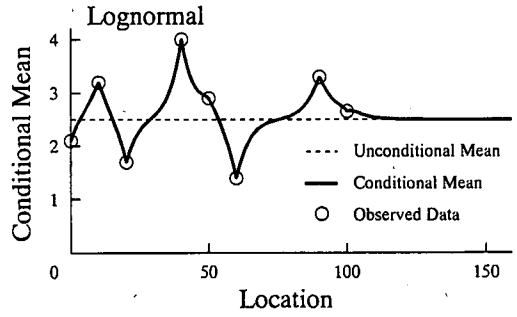
分布形の違いによらず, シミュレーション結果は理論解とほぼ一致している。このことから文献4)で提案した理論の妥当性が確認できた。なお, 観測点から離れた内外挿点の条件付サンプル分散値には若干の変動が見られるが, この点はサンプル場の回数を増やせば解決される問題である。

内外挿点の条件付平均値は無条件平均値に近づくように求められている。観測点から離れた遠方場では観測値に左右されにくくなるため, 最適推定値は無条件平均値と一致する。推定点と観測点一致すると, 条件付平均値すなわち最適推定値は観測値と完全に一致する。若干の差異は見られるものの, 分布形によらず推定値はほとんど変わらない。このことは, 確率場の分布特性に応じた推定理論式を用いることなく, 計算が容易ということから, これまで伝統的に採用されてきた条件付正規確率場の推定式(式(42))によって近似的に条件付非正規確率場を取り扱うことができることを意味しない。

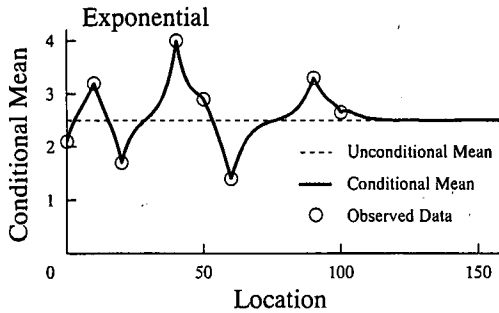
一方, 条件付分散では分布形によって差が生じている。正規分布, 一様分布のように対称分布の場合, 条件付分散は無条件分散を超えることはない。他の分布形(対数正規, 指数, レイリー, グンベル分布)では無条件平均値よりも大きな観測値に依存して, 条件付分散が無条件分散よりも大きくなることもある。観測値にも左右されるが, 一般に密な観測点配置であるほど, その内挿点での条件付分散は小さい。これは直感的な感覚と一致するものである。



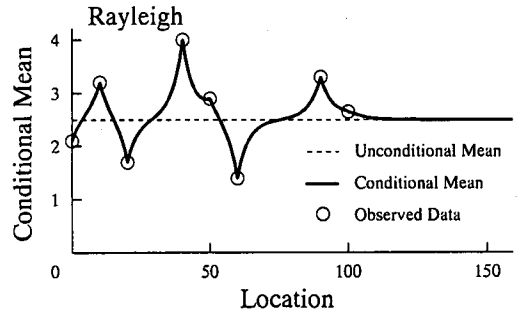
(a)Gaussian distribution



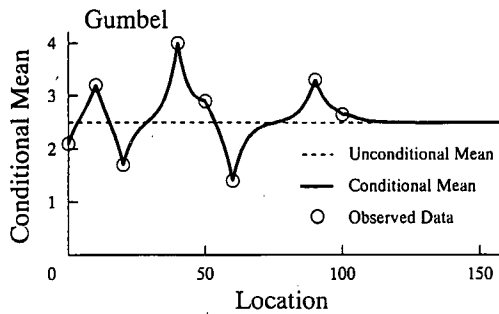
(b)Lognormal distribution



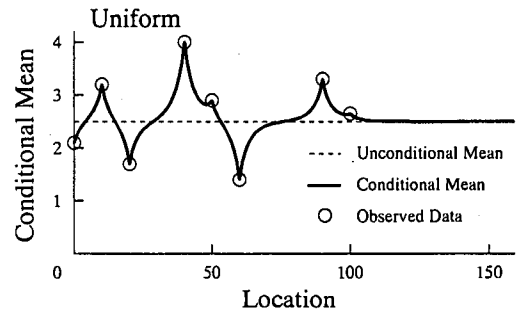
(c)Exponential distribution



(d)Rayleigh distribution



(e)Gumbel distribution



(f)Uniform distribution

Fig.1 Theoretical solution to conditional mean

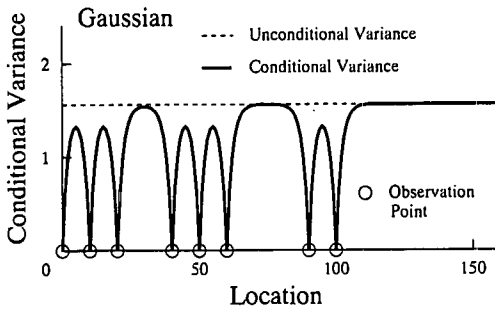
ここでは式(41)の推定誤差分散の数値結果を示していない。しかしながら同式の多重積分を解析的に評価可能な条件付対数正規確率場については、条件付分散との相違点などに関し、文献4)で既に議論しているので、それを参照されたい。

(3) 最大エントロピー推定法の結果

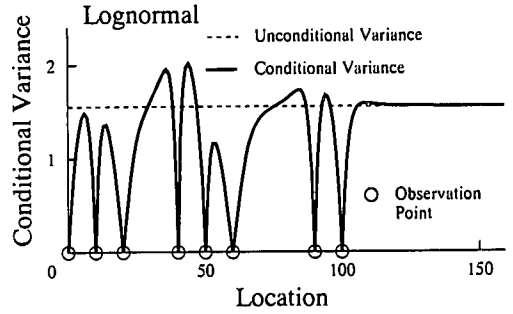
最大エントロピー推定値は式(42), (49), (59), (65), (71), (76)によって計算する。式(65), (71)の実超越方程式の近似根を求めるためには初期値を与えなければならない。そのためここでは条件付正規確率場の最大エントロピー推定値(式(42))を初期値とした。各分布

形の最大エントロピー推定値はFig.5のようになる。

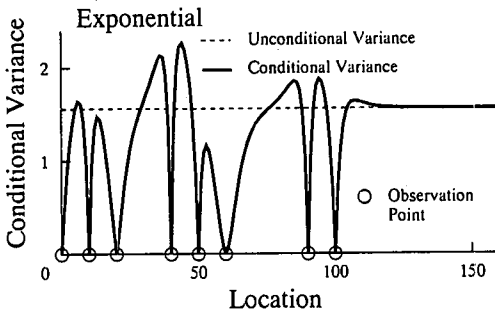
最大エントロピー推定値は式(24)のエントロピーを最小にする解であるが、これは条件付確率密度関数の最大値を与えるモードにほかならない。従って内外挿地点の推定値は、一様分布の場合を除き、無条件確率密度関数のモードに近づくように補間されている。正規確率場ではモードと条件付平均値が一致するので、Fig.1とFig.5の結果に差が見られない。無条件非正規確率場のモードは無条件平均値よりも小さいので、対数正規、指数、レイリー、グンベルの各分布に対する最大エントロピー推定値は滑らかではなく、偏って推定されていることがわかる。



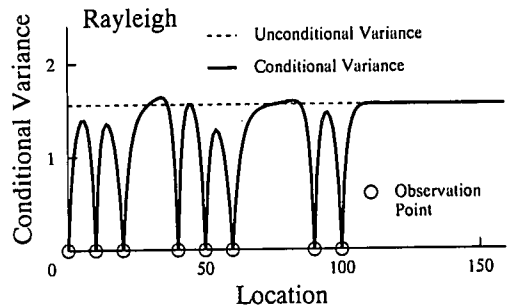
(a) Gaussian distribution



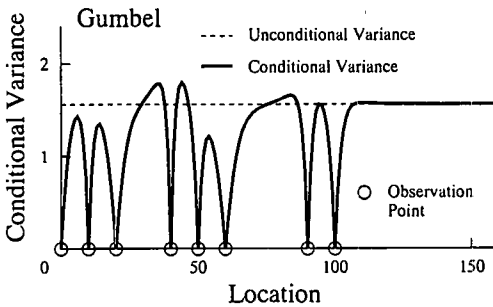
(b) Lognormal distribution



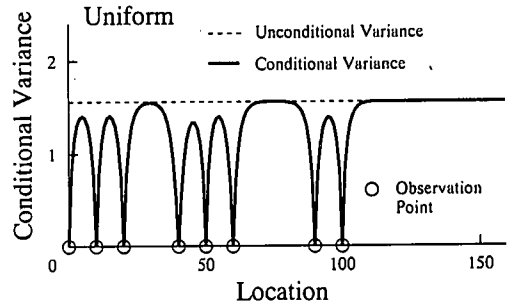
(c) Exponential distribution



(d) Rayleigh distribution



(e) Gumbel distribution



(f) Uniform distribution

Fig.2 Theoretical solution to conditional variance

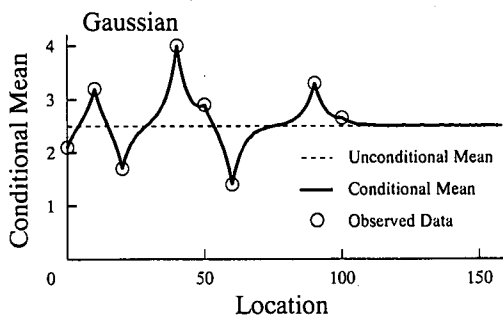
一様分布の最大エントロピー推定値はモードが特定化されないため、観測値の影響が小さくなると、中央値である無条件平均値よりも大きい観測値近傍の推定点では境界値の b_n に、無条件平均値よりも小さい観測値近傍では境界値の a_n になるように推定値が求められている。観測点から離れた遠方場の最大エントロピー推定値は無条件平均値に至っているが、これは本論の定義に従ったからにはかならない。

式(24)のエントロピーの最小化は一様誤差の最小化と等価であり、この規範に基づいて得られたのがFig.5である。一方、Fig.1の結果は推定誤差分散の最小化に基づいて求められたものである。従ってFig.1が不偏推定

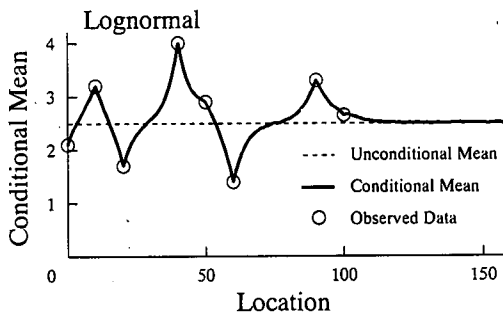
値に対応するのに対し、Fig.5はMAP推定値となる。このことから対象とする物理的問題、推定誤差の最小化問題に応じて、不偏推定値すなわちKrigingの最適推定値か、最大エントロピー推定値すなわちMAP推定値のどちらかを選択する必要がある。

Fig.6は式(45), (54), (63), (69)(ただし最初の無限区間積分項は無視する), (74), (79)に基づいて計算された条件付エントロピー $-\varepsilon(X_n|C)$ を示したものである。既に述べたように、このエントロピーは負の値にもなり得る。特に観測点では $-\infty$ の値となる。そこでここでは図の体裁上正のエントロピーのみを示す。

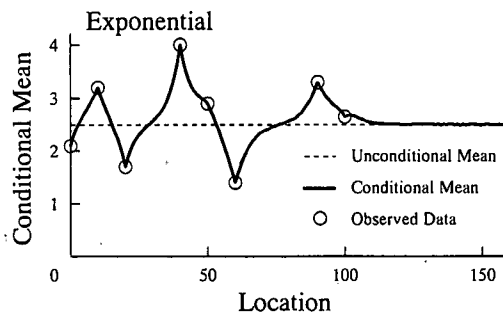
正規分布、一様分布の条件付エントロピーは無条件



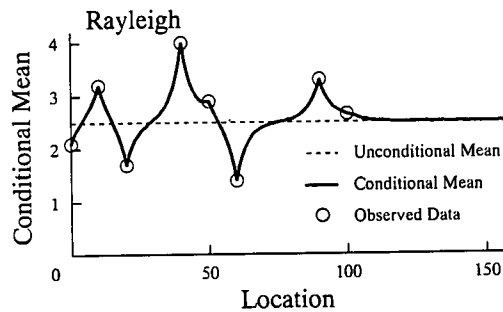
(a)Gaussian distribution



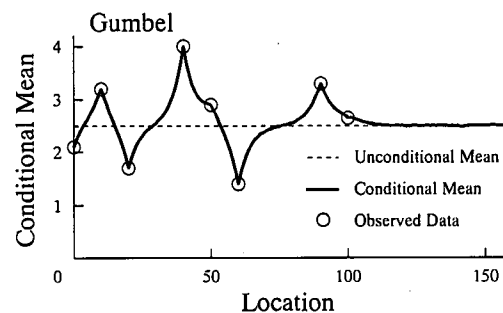
(b)Lognormal distribution



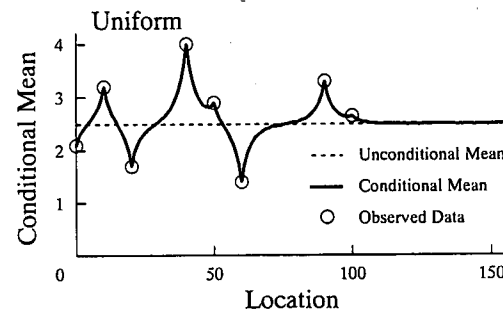
(c)Exponential distribution



(d)Rayleigh distribution



(e)Gumbel distribution



(f)Uniform distribution

Fig.3 Conditional mean obtained by conditional simulation technique

エントロピー $H(X_n)$ (式(44), (78)) よりも大きくなることはない。一方他の非正規確率場の条件付エントロピーを見ると、無条件平均値よりも大きい観測値の影響を受け、その観測点近傍の値は無条件平均エントロピー (式(53), (62), (68), (73)) よりも大きくなっている。観測点から離れた遠方場の条件付エントロピーは無条件平均エントロピーに漸近する。観測点では観測値が確定的に与えられるため、条件付エントロピーは $-\infty$ の値となる。このような条件付エントロピーの特徴は Fig.2 の条件付分散の性質と同じである。

式(46), (55), (64), (70), (75), (80) によって条件付平均エントロピー $H(X_n|C)$ を計算すると、Fig.7 の実線

のようになる。Fig.6 と同様な理由により、正の部分のみ図化した。図中には無条件平均エントロピー $H(X_n)$ を破線で示している。破線の値から実線の値を引くと、平均相互情報量 $I(X_n; C)$ (式(48) 参照) が得られる。このことは Fig.7 から明らかに予想されることなので、その図は割愛する。結果として、1) 分布形によらず条件付平均エントロピーあるいは平均相互情報量の推定点ごとの大小関係は同じであること、2) これらは観測値に依存せず、観測点位置のみに左右されること、3) 条件付平均エントロピーは無条件平均エントロピーに等しいかそれよりも小さく、従って $I(X_n; C) \geq 0$ が成り立つこと、4) 観測点の条件付平均エントロピーは $-\infty$

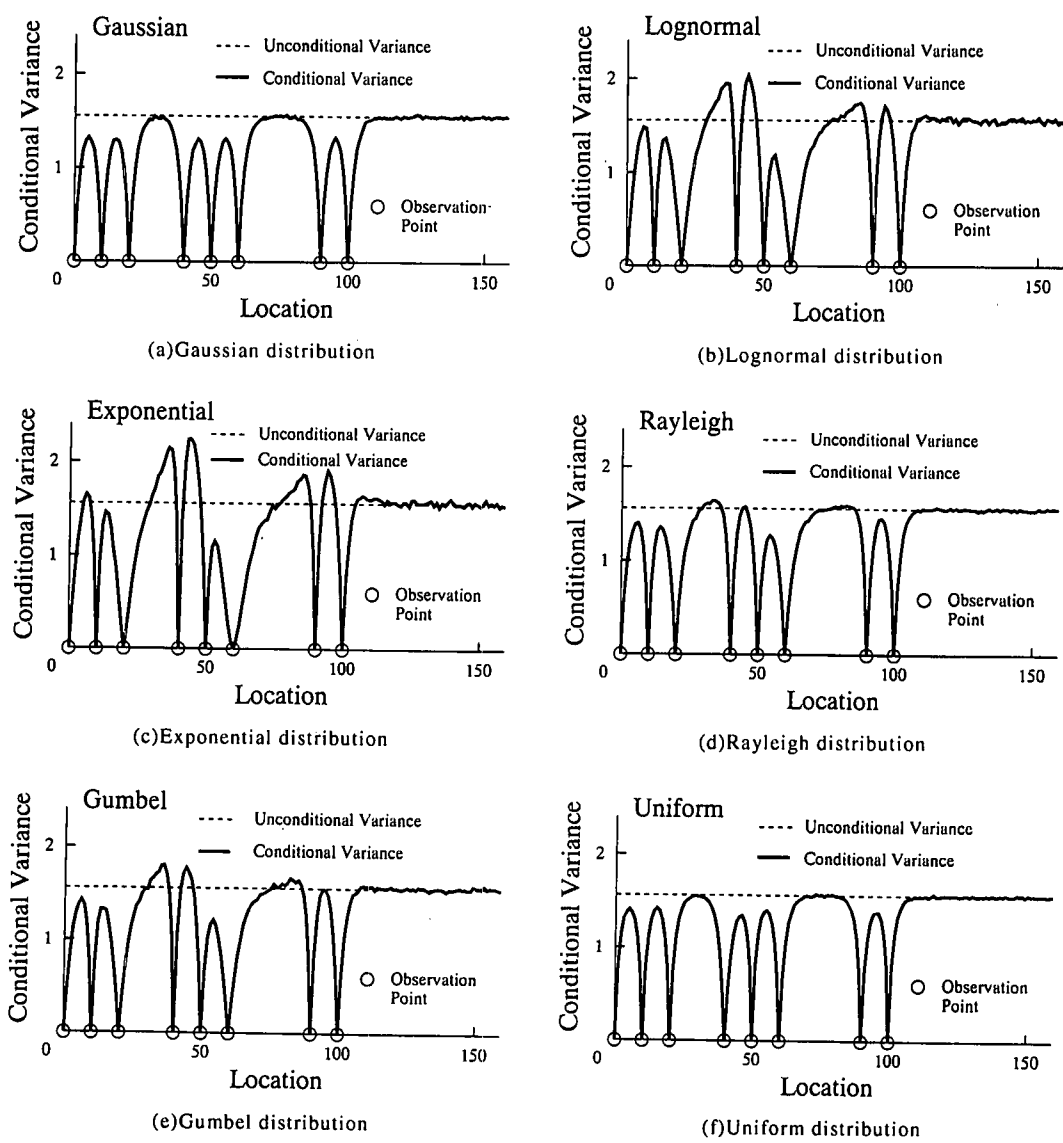


Fig.4 Conditional variance obtained by conditional simulation technique

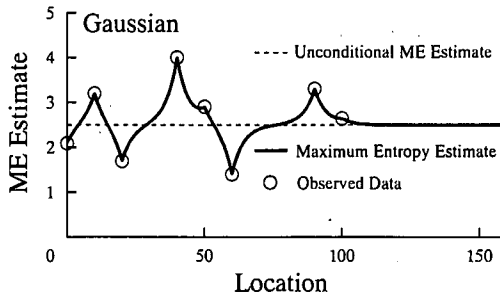
になること、5) 観測値の影響が小さくなると(例えば遠方場において)、条件付平均エントロピーは無条件平均エントロピーに一致することが見いだされる。3)、4)についてはShannonの不等式そのものである。

条件付平均エントロピーは推定誤差を直接表わす指標ではないが、その特徴は推定誤差分散と同一である。式(41)に示す推定誤差分散の評価式は多重積分を含んでいる。従って分布形によっては解析的にこれを求めることができず、数値積分を行う必要がある。実際には多重積分を実施することが大いに困難なので、このような場合には推定誤差分散に代わって条件付平均エントロピーを用いるメリットがあろう。

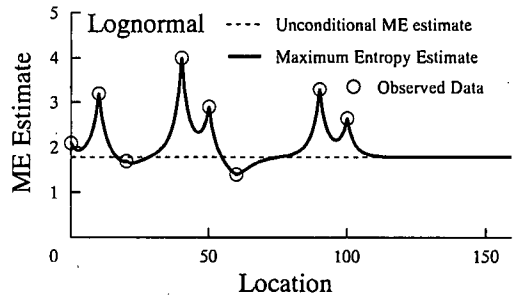
8. 結論

Kriging法は、観測値が与えられたとき、不偏性を満たした上で未観測点における真値と推定値の誤差を最小とするように、各点での最適推定値と推定誤差分散を求めようとするものである。最も基本的なKriging法はSimple Krigingと呼ばれ、空間場の確率特性が事前に既知として設定される。現在ではこの概念を拡張して種々のKriging手法が考案されている¹⁾。

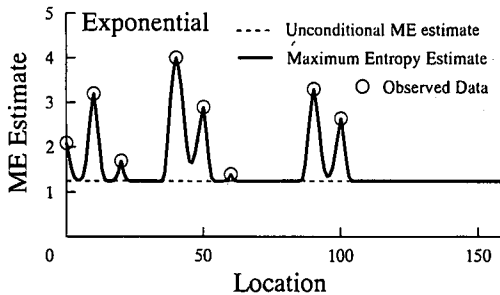
Kriging手法は条件付確率場の推定理論の一つであるが、不偏性と推定誤差分散の最小化を規範として、最適推定値と推定誤差分散が得られる。本論文ではこの規



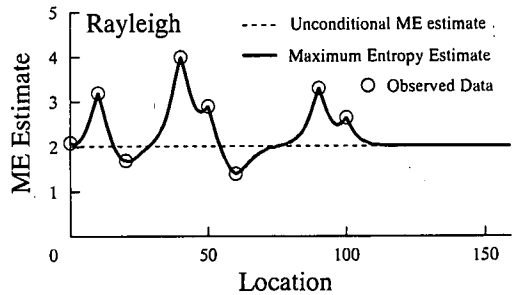
(a)Gaussian distribution



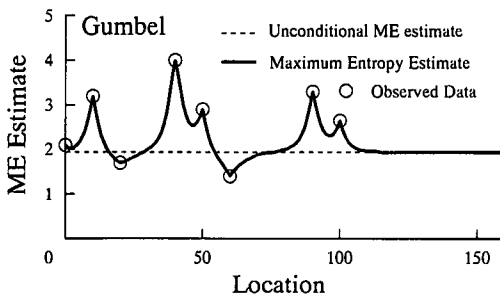
(b)Lognormal distribution



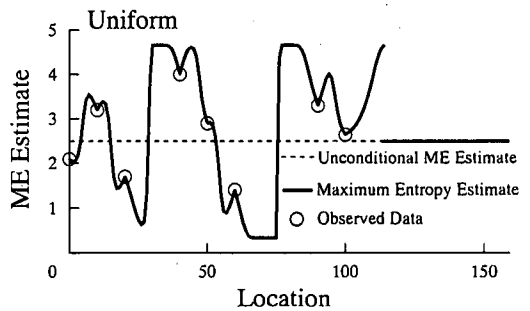
(c)Exponential distribution



(d)Rayleigh distribution



(e)Gumbel distribution



(f)Uniform distribution

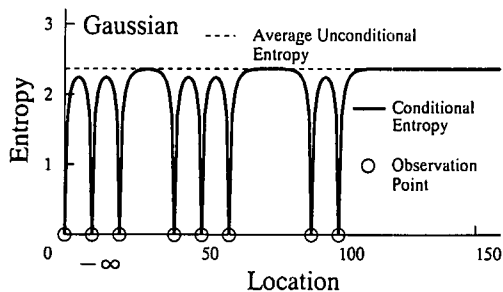
Fig.5 Effect of differences in distribution type on maximum entropy estimate

範を用いず式(24)のエントロピーの最小化を規範として、条件付非正規確率場における最大エントロピー推定値を求める方法を提案し、Kriging法の結果と比較検討した。また観測情報に基づいて事前に設定した確率場の更新程度を推定誤差分散ではなく、条件付平均エントロピーによって評価した。理論ならびに数値解析により、提案手法とKriging法の相違点・類似点を明らかにするとともに、条件付非正規確率場の推定理論に関して新たな知見を得た。以下に主な成果を示す。

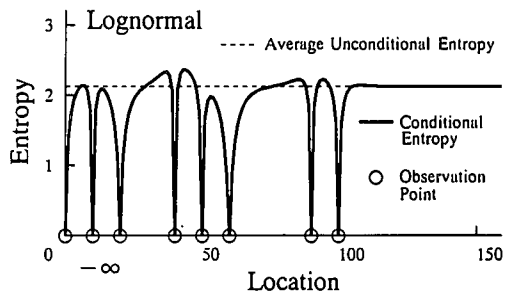
- 1) 星谷・野田・稲田⁴⁾によって提案された非正規確率場の条件付確率密度関数から未観測点のエントロピーを定義し、その最小値を与える解を最適推

定値と考えた。この解は文献(12)、(13)と対応させて最大エントロピー推定値と名付けたが、MAP推定値と一致する。

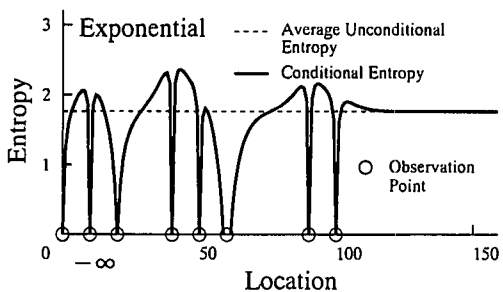
- 2) 未観測点における無条件確率場のエントロピー(無条件平均エントロピー)、観測値が得られたときの条件付確率場における条件付エントロピー、それを観測点の確率分布で平均した条件付平均エントロピーの一般式を誘導した。
- 3) 最大エントロピー推定値、条件付エントロピーは観測値に依存するが、条件付平均エントロピーは観測位置のみに依存し、観測値に無関係である。ただし正規確率場における条件付エントロピーと条



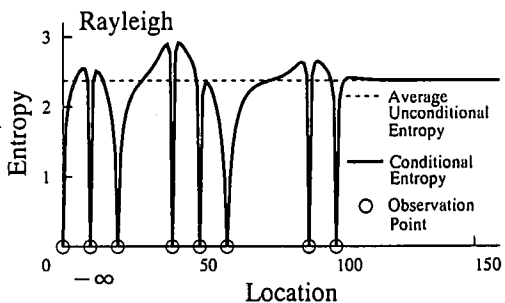
(a) Gaussian distribution



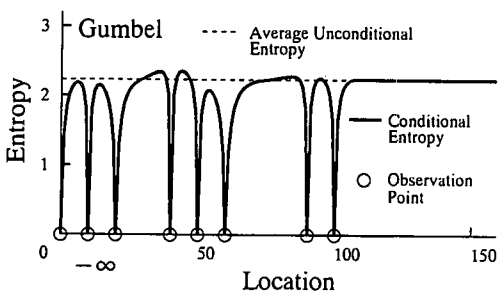
(b) Lognormal distribution



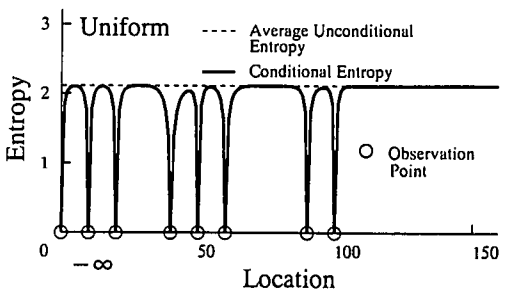
(c) Exponential distribution



(d) Rayleigh distribution



(e) Gumbel distribution



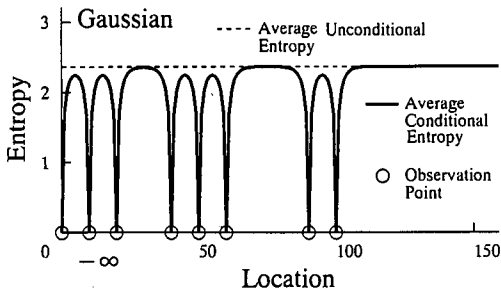
(f) Uniform distribution

Fig.6 Effect of differences in distribution type on conditional entropy with value $-\infty$ at observation point

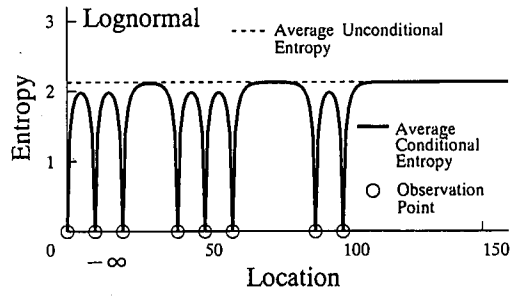
条件付平均エントロピーは同一値となり、観測値に依存しない。

- 4) 条件付平均エントロピーすなわち事後確率分布に関する平均エントロピーは無条件平均エントロピーすなわち事前確率分布のエントロピーよりも大きくはならない。これより先験知識などで与えた不確かさの量は観測情報に基づき減少し、決して増大することはない。このことはShannonの不等式としてよく知られたことではあるが、条件付確率場の推定問題においても成立する。ただし条件付エントロピーは観測値に依存するので、無条件平均エントロピーより大きくなることもある。

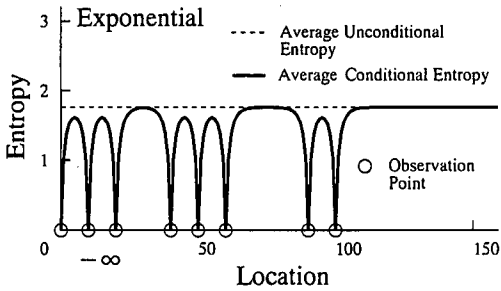
- 5) 3), 4)の性質は不偏推定・最小誤差分散規範に基づく星谷らの推定法の結果と同一である。すなわちKriging法の最適推定値は最大エントロピー推定値に、条件付分散は条件付エントロピーに、推定誤差分散は条件付平均エントロピーに対応した性質を有する。
- 6) 提案手法を正規確率場ならびに5つの非正規確率場(対数正規、指数、レイリー、グンベル、一様分布)に適用して推定理論式を導いた。Kriging手法ならびに最大エントロピー推定法の最適推定値は正規確率場で同一解を与え、推定誤差分散の対数値は条件付平均エントロピーになる。このこと



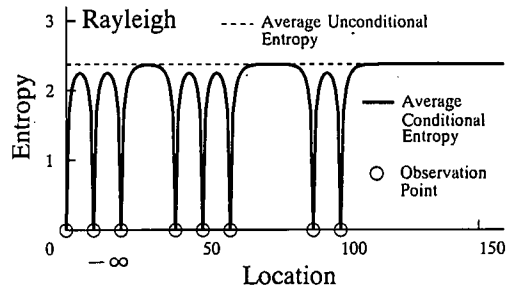
(a)Gaussian distribution



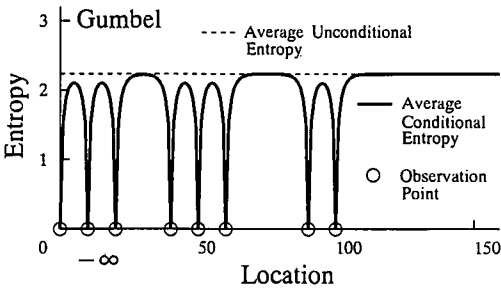
(b)Lognormal distribution



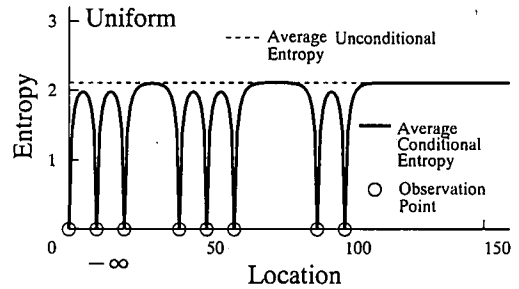
(c)Exponential distribution



(d)Rayleigh distribution



(e)Gumbel distribution



(f)Uniform distribution

Fig.7 Effect of differences in distribution type on average conditional entropy with value $-\infty$ at observation point

より、両手法は等価であると言える。一方両手法は対数正規分布において異なる結果を与えるが、密接に関連している。最大エントロピー推定値はKrigingの最適推定値よりも小さく、滑らかでない補間値を与えることがわかった。

- 7) 観測点から離れた遠方場の最大エントロピー推定値は観測値の影響を受けにくい。対数正規分布、指数分布、レイリー分布、グンベル分布に従う条件付非正規確率場によると、観測点から遠方の未観測点の推定値は無条件確率場の周辺分布のモードとなり、無条件平均値よりも小さく求められる。これは提案手法が不偏性を念頭において定式化され

ていないためである。なお正規分布は対称性を有するため、遠方場の最大エントロピー推定値は無条件平均値と一致する。一方、一様分布ではモードが特定できないため、境界値が最適推定値となることがある。

- 8) 未観測点すなわち推定点と観測点が一一致すると、最大エントロピー推定値は観測値と完全に一致し、その結果条件付エントロピー、条件付平均エントロピーは $-\infty$ となる。この意味で本論文で提案した最大エントロピー推定法は条件付推定問題の解法と言える。
- 9) 最大エントロピー推定量に対応した推定誤差分散

は星谷・野田・稲田の厳密解法によって算定可能である。しかしながら解析的に積分可能な形で与えられる場合を除き、その評価には多重積分が必要となり、その数値処理は必ずしも容易ではない。一方、本論文で提案した条件付平均エントロピーは解析的に求められることが多い。これは推定誤差を厳密に表す指標ではないが、不確実性の評価に当たっては推定誤差分散に比べて有用な指標となりうる。

- 10) 無条件平均エントロピーから条件付平均エントロピーを引くと、平均相互情報量が得られる。平均相互情報量は推定誤差分散を陽に与えるものではない。しかしながら、これは事前情報に基づいたモデルの不確実性が観測情報に基づきどの程度低減できるかを計量する定量的な尺度であり、条件付確率場の更新程度を評価する上で便利な指標である。

以上に示すように、非正規分布の最大エントロピー推定量あるいは最大エントロピー推定値には偏りが見られることがわかった。一方で推定精度を厳密に表す指標ではないが、観測に伴う更新程度については平均相互情報量を用いて比較的容易に評価できる。このように最大エントロピー推定法には一長一短がある。従って、一般的には観測値を用いて内外挿面を求めるのであれば、不偏推定・最小誤差分散規範に基づく星谷・野田・稲田の推定理論を、観測によって事前確率場の更新程度を分析する場合には本論文で提案した最大エントロピー推定法を用いることが望ましい。なお解析的に推定誤差分散が求められる場合には、この推定誤差分散を用いることによって不確実性の減少度合を計量化できるので、Kriging法が最大エントロピー推定法よりも優れていると考えられる。

参考文献

- 1) Cressie, N. A. C.: *Statistics for Spatial Data*, Revised Edition, John Wiley & Sons, New York, 1993.
- 2) 星谷 勝: 条件付確率場のシミュレーション理論, 土木学会論文集, No.459/I-22, pp.113-118, 1993年1月.
- 3) Shinozuka, M. and Zhang, R.: Equivalence between kriging and CPDF methods for conditional simulation, *Journal of Engineering Mechanics*, ASCE, Vol.122, No.6, pp.530-538, June 1996.
- 4) 星谷 勝, 野田 茂, 稲田 裕: 観測情報に基づく条件付非正規確率場の推定理論の誘導, 土木学会論文集, No.571/I-40, pp.83-95, 1997年7月.
- 5) 長谷川 淳, 井上光洋共訳 (Shannon, C. E. and Weaver, W. 著): コミュニケーションの数学的理論 - 情報理論の基礎 -, 明治図書, 1969年5月.
- 6) 池原止戈夫, 彌永昌吉, 室賀三郎, 戸田 巖共訳 (Wiener, N. 著): サイバネティクス, 第2版, 岩波書店, 1991年5月.
- 7) 有本 卓: 現代情報理論, 電子情報通信学会, 1988年1月.
- 8) 本城勇介, 工藤暢章: 情報エントロピーによる逆解析のための観測計画の評価方法に関する基礎的考察, 土木学会論文集, No.589/III-42, pp.321-333, 1998年3月.
- 9) 吉田郁政: 未知パラメータ空間の基底変換を用いた逆解析, 土木学会論文集, No.577/I-41, pp.205-215, 1997年10月.
- 10) Ditlevsen, O. and Madsen, H. O.: *Structural Reliability Methods*, John Wiley & Sons, Chichester, 1996.
- 11) Liu, P.-L. and Kiureghian, A. D.: Multivariate distribution models with prescribed marginals and covariances, *Probabilistic Engineering Mechanics*, Vol.1, No.2, pp.105-112, June 1986.
- 12) Lee, Y.-M. and Ellis, J. H.: On the equivalence of kriging and maximum entropy estimators? *Mathematical Geology*, Vol.29, No.1, pp.131-152, January 1997.
- 13) Christakos, G. and Li, X.: Bayesian maximum entropy analysis and mapping: A farewell to kriging estimator? *Mathematical Geology*, Vol.30, No.4, pp.435-462, May 1998.
- 14) 富士通: FACOM FORTRAN SSL II 使用手引書 (科学用サブライブラリ), 1970年12月.

(1998. 7. 28 受付)

MAXIMUM ENTROPY ESTIMATOR OF CONDITIONAL STOCHASTIC FIELDS

Shigeru NODA and Masaru HOSHIYA

This paper proposes a theory for evaluating the maximum entropy estimators and their associated entropy indices for conditional non-Gaussian translation stochastic fields when observation is made at some discrete points. Through analytical development and numerical examples of stochastic fields with the six types of distribution, kriging and maximum entropy techniques for spatial estimation are also compared. It was found that: 1) the maximum entropy estimate and conditional entropy are dependent on the values of observed data; and 2) the average conditional entropy has an independence on the observations, and is not larger than the average unconditional entropy.