

参 考 文 献

Yoshida, N. (1994). Malliavin calculus and asymptotic expansion for martingales, Research Memo., No. 504, The Institute of Statistical Mathematics, Tokyo.

知的統計モデルについて

松 縄 規

統計科学に於ける各種モデルの位置づけと意味を考察した。モデルとして基礎モデル，発展モデル，知的モデルを提案した。基礎モデルは広い意味でのデータに基づいて構築される。この際筆者がこれまでに考えてきた観測対象と観測機構の間の統計的不確定性を考慮しそのことと関連する統計的基礎方程式を利用して基礎モデル P を構築する立場を取った。

統計モデルの一つの理想は数理に裏打ちされた統計基礎構造，その構造変化を検知する機構，その変動を制御する機構，それらに加えてモデル自身が学習する機能を持つことである。これらの共通部分の機能を持つモデルは一種の知的モデルと見做し得るものでありそれについて他のモデルとの関連と位置づけを述べた。このモデルほどには水準は望めないが部分的にそれに近い機能をもつものとして発展モデルを考えた。

基礎モデル P が定まっている時，その平均等の統計的状态が断続的に変化すると P はどう変化するだろうか？ これに対し P から出発し各段階で最適なモデルの列 $\{Q_k\}$, ($k=1, 2, \dots$) を構成，第 m 段階で，予め設定した許容誤差の範囲内で，未知の発展モデル Q に原理的に到達出来る。 Q は K-L 情報量が $I(Q; Q_m) \rightarrow 0$ ($m \rightarrow \infty$) となる様に探索をし実現される： $F_k = \{Q_i; i=1, 2, \dots, k\}$, ($k=1, 2, \dots$) を独立な k 個の平均条件を満たす確率分布族の縮小列とする。ルジャンドル変換により K-L 情報量間の情報収支が成立つ： $I(Q; Q_{k-1}) = I(Q; Q_k) + I(Q_k; Q_{k-1})$ ，ここに $Q_0 \equiv P$, $Q_k \in F_k$ ($k=1, 2, \dots$)。右辺第1項は分布システム $\{Q; Q_k\}$ の内部生成情報量を，第2項は第 k 段更新時の基礎モデル分布 Q_{k-1} から更新モデル Q_k への輸送情報量を表わす。上述の更新に対応し内部生成情報量の単調減少列を構成出来る。従って $I(Q; Q_m) \rightarrow 0$ ($m \rightarrow \infty$) を得る。輸送情報量についても $I(Q_m; Q_{m-1}) \rightarrow 0$ ($m \rightarrow \infty$) となるから， $P \rightarrow Q_1 \rightarrow Q_2 \rightarrow \dots \rightarrow Q_m$ の様にモデルを逐次更新して，十分大きな m に対し $Q \cong Q_m$ と出来る。

空間における smoothness prior の使い道について

尾 形 良 彦

地震の震源位置は通常，各観測地点で計測された地震波の到達時刻のデータを使って最小自乗法によって推定されるが，正確に決定するためには震央の周りに十分密な観測点が必要である。現在，世界には1500に昇る常時観測地点がある。数多くの観測地点の到達時刻データを集めようとするほど震源決定までには長い時間がかかるが，こうして決められた震央分布は精密で，たとえば地球の表面がいくつかのプレートに分割されていることがくっきりと示される。

長野県松代にある気象庁地震観測所の群列地震観測システム (MSAS) は半径5 km ほどの円周と中心の7地点にほぼ等間隔に置かれた地震計からテレメータによって同時に地震波データを集約するもので，地震波が観測されれば，これらによってだけでも広域的な震央を決める

ことが可能である。しかし松代周辺 100 km 以内を除けば、決められた震央は大分実際と異なった分布をしている。この原因としては、P 波や S 波の不明瞭さによる到着時刻の読み取り誤差に基づくランダム (unbiased) なものと、地震波の伝播経路の速度構造の不均質性に基づく系統的 (biased) なものが考えられる。いま MSAS によって推定された震央の位置を (x_p, y_p) 、真の震央の位置を (u_p, v_p) とし、前者のランダム誤差を ϵ_p, η_p 、後者の系統的な偏差を関数 $f(x_p, y_p), g(x_p, y_p)$ で表すと、これらの関係は次の式

$$\begin{aligned} u_p &= f(x_p, y_p) + \epsilon_p \\ v_p &= g(x_p, y_p) + \eta_p, \quad p=1, 2, \dots, N \end{aligned}$$

によって示される。

本報告の目標は偏差関数 $f(x, y), g(x, y)$ がある程度滑らかであると仮定して、これらを求めて、将来の MSAS 震央から震央の補正写像

$$\varphi: (x, y) \mapsto (u, v) = (f(x, y), g(x, y))$$

によって実際に近い震央を予測することである。真の震央の位置 (u_p, v_p) として、多観測所の地震波到達時刻データを用いて決定された気象庁震源カタログ (日本近辺) や米国地質調査所の Earthquake Data Report (世界) の震央データを使用する。他方、たとえばバイアス関数 f は二次元 B スプライン曲面 $f = f(x, y) = \sum_{i,j} \theta_{i,j} B_i(x) B_j(y)$ を用いて滑らかさに関する制約

$$\Phi_1(f) = \iint \left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 dx dy, \quad \Phi_2(f) = \iint \left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 dx dy$$

のもとでの最小自乗平滑化

$$\sum_{p=1}^N \{u_p - f(x_p, y_p)\}^2 + w_1 \Phi_1(f) + w_2 \Phi_2(f)$$

を行う。このとき制約もパラメータの二次形式になっているから、事前分布が多変量正規であり、従って事後分布も多変量正規である。かくしてベイズ尤度が計算可能となり、最適な重み w_1 および w_2 は ABIC を最小化するものとして得る事ができる。同様にして関数 g についても最適解を求めることが出来る。 f, g 各々のパラメータ $\theta_{i,j}$ の数が多ければ多い程 ABIC が減少するが、約 100×100 個までには最小 ABIC が収束しているようである。

上記の方法で 1984~88 年間のデータによって推定された補正写像 $\varphi = (f, g)$ を用いて 1984~88 年間の MSAS 震央から世界及び日本の震央分布を求めたところ良好な結果が得られた。また残差解析の結果、 $\{\epsilon_p\}$ と $\{\eta_p\}$ は無相関であること、そして MSAS の震央の角度 $\psi = \tan^{-1}(y/x)$ に関する推定誤差が震央までの距離 $\Delta = \sqrt{x^2 + y^2}$ に関する推定誤差より大きく影響していることが示される。

同一指数の regularly varying tail をもつ分布の Mellin-Stieltjes convolution

志村隆彰

X, Y をそれぞれ分布 μ, ν に従う正值の独立確率変数とする。このとき、これらの独立積 XY の分布を μ と ν の Mellin-Stieltjes convolution と呼び、 $\mu \circ \nu$ で表す。独立同分布に従う確率変数列に関する極限定理と関連する分布族である $D(\alpha)$ ($\alpha \geq 0$) (正の台をもち、tail が指数 $-\alpha$ の正則変動をする分布族) に属する分布の Mellin-Stieltjes convolution について、以下