

先ず、多変量極値分布の応用として、大阪と神戸の年間最高気温データを解析した。データは2次元の極値分布に従うと仮定する。各周辺分布に一般化極値分布族をあてはめそのパラメータを推定する。この推定値を利用して、各データが平均1の指数分布からのデータとみなせるように変換する。この2次元データを Pickands の2次元指数分布からのデータとして、その従属関数を Pickands のノンパラメトリック法と Tawn (1988) のパラメトリック法で求めた。求めた従属関数によって大阪と神戸の年間最高気温間の関係を見た。

次に、多変量極値統計量の成分が漸近独立や完全従属になるための必要十分条件について調べた。周辺独立や完全従属になる多変量極値分布は1点の性質で特徴づけられ、この性質は弱収束の場合にも保存される (Takahashi (1993))。このことを用いて、多変量極値統計量の成分が漸近独立になるための簡単に使いやすい必要十分条件を分布関数 F や F の dependent 関数を用いて与えた。同様のことを完全従属の場合にも示した。

参 考 文 献

- Takahashi, R. (1993). Asymptotic independence and perfect dependence of vector components of multivariate extreme statistics, *Statist. Probab. Lett.* (to appear).
- Tawn, J.A. (1988). Bivariate extreme value theory: models and estimation, *Biometrika*, **75**, 397-415.

統計教育・情報センター

相同性解析に基づく全生物リボソーム蛋白質の編集

橋 本 哲 男

リボソームは地球上の全ての生物に共通に存在する細胞内小器官であり、数種の核酸 (RNA) 分子と 50~80 種にのぼる蛋白質分子の複合体 (超分子) として構成されている。リボソームは細胞内で必要な全ての蛋白質分子を合成する場、すなわち、遺伝情報翻訳の場であり、生命現象を維持するうえで最も基本的な役割を担っている。リボソームを構成する分子 (核酸, 蛋白質) に関する構造データは、近年の遺伝子工学的手法の発展に伴い、広範な生物種にわたり数多く蓄積されてきている。これらのデータは、地球上に存在する生物の進化の過程を解明するうえで貴重な情報を含んでいると考えられる。現在、リボソームを構成する蛋白質の一次構造データに関しては、50 種類以上の生物にわたり約 1,000 種が報告され、国際的な遺伝情報データベースに登録されている。三大生物群の中で真正細菌群に属する大腸菌に関しては、構成蛋白質 54 種の全一次構造データが既知であるが、我々ヒトを含む真核生物群、および古細菌 (後生細菌) 群に属する生物種に関しては、全蛋白質種の一次構造データは明らかとなっていない。

我々は、リボソーム蛋白質を対象とした分子進化的解析を行う際の基礎資料を得ることを目的として、約 1,000 種の登録蛋白質のなかから大腸菌全 54 種のそれぞれに相同と考えられる蛋白質を「相同性解析」のアルゴリズムに基づいて同定し、大腸菌の命名法に従ってこれらを分類、統合した。さらに、アミノ酸の物理化学的性質から推定される高次構造の類似度を指標として相同蛋白質同士の相似度マトリクスを構成し、これをもとに全 54 種のアライメントを構築した。その結果、以下に示すことが明らかとなった (Otaka et al. (1993))。 (1) 三大生物群を通して存在が認められる蛋白質種が少なからざる数にのぼる、 (2) 多くのアライメントは、挿入、欠失、分子伸長等の変化を生じているものの、ほぼ一対一対応のパターンを示す、すなわち、分子内転座 (大腸菌 L7/L12 相当)、分子内二重転座 (S6 相当)、二分子融合 (L10 相当) などの大きな変異は稀である、 (3) 大幅な挿入、欠失、分子伸長などの変化は、三大生物群同士の間で見いだされる、逆に、同一群の中では、アミノ酸置換が主たる変異の要因となっている。

参 考 文 献

Otaka, E., Hashimoto, T. and Mizuta, K. (1993). The ribosomal proteins: I. An introduction to a compilation of the protein species equivalents from various organisms by a universal code system, *Protein Sequences and Data Analysis* (in press).

Fokker-Planck 方程式の一般化と繰り込み

岡 崎 卓

1. はじめに

風波を受けつつ航行する船舶の傾角や、外乱の混入を避け難い電気回路の出力は不規則な変動を呈する。外乱の統計を知って変動の確率分布を予測することは、船体や電気回路を設計するうえで欠かすことができない。Fokker-Planck (FP) 方程式はこの確率密度を定める方程式として多用されているが、その厳密な成立は外乱が正規白色ノイズである場合に限られる。

従って、任意の外乱に対して正しく確率密度を定めるよう FP 方程式を一般化することが望ましい。以下では、互いに関連しつつ発展する複数の変動を対象に、Generalized Fokker-Planck (GFP) 方程式を導き、さらに簡約する方法のあらましを述べる。

2. GFP 方程式の導出

確率変数 $U = \{U_i, i=1, \dots, n\}$ と正規白色ノイズ $\nu(t)$, $\langle \nu(t)\nu(s) \rangle = \delta(t-s)$ で駆動される外乱 W の時間発展を記述する確率微分方程式

$$\begin{aligned} \frac{d}{dt} U_i &= M_i(U) + \mu_i(U, W), & i=1, \dots, n \\ \frac{d}{dt} W &= N(W) + \sigma^2 \nu(t) \end{aligned}$$

が与えられれば、 U と W の結合確率密度 $\rho(U, W, t)$ を決定する方程式は $n+1$ 箇の変数に関する FP 方程式として容易に得られる。この方程式から外乱 W を除いた主変数 U のみの確率密度 $f(u, t)$ を求めることが FP 方程式の一般化にはほかならない。

さて、 (U, W) の任意の位相関数 $X(U, W)$ を $\prod_{i=1}^n \delta(U_i - u_i)$ の線形空間に写す射影子を外乱 W の密度関数を介して組み立てれば、統計力学の手法を用いて、 $f(u, t)$ を定める f のみを含む GFP 方程式を形式的に書き下すことができる。

3. GFP 方程式の簡約

上記 GFP 方程式は、作用素で形式的に表現され具体的には計算不可能な係数を含み、そのままでは実用性に乏しい。そこで微少展開パラメータ ε により $\frac{d}{dt} U = M(U) + \varepsilon \mu(U, W)$ と書替え、 f および拡散係数を ε で摂動展開する。各次数毎に f の形式解を求めた後、再び総合して $\frac{\partial}{\partial t} f$ を復元すると、GFP 方程式はもはや計算困難な項を含まない。

こうして得られる簡約 GFP 方程式は拡散方程式に類似の構造をもち、その拡散係数は例えば

$$\begin{aligned} \lambda_{ij}(u) &= \int_0^t ds B_{ij}(u, s) \cdot R_{ij}(u, -s) \\ B_{ij}(u, s) &= \langle \mu_i(U(-s), W) |_{U=u} \mu_j(u, W(s)) \rangle \end{aligned}$$