

多次元時系列の大域的相関分析法

3 P-5

A global correlation analysis of multidimensional time series

加藤 崇利† 金子 茂宏† 佐藤 貴志† 中島 博史† 吉川 昭† 小原 啓義†
 Takatoshi KATO Shigehiro KANEKO Takashi SATO Hiroshi NAKAJIMA Sho KIKKAWA Hiroyoshi OHARA

1 はじめに

多次元時系列の解析において、観測される複数の時系列（以下では時系列束と呼ぶ）の相関関係を定量的に評価することは重要である。しかし従来の相関分析法で用いられる自己および相互相関関数は時系列中の2点間の相関値に基づくものであり、多次元時系列全体の特徴を捉えるのに必ずしも有用ではない。そこで本報告では、時系列束の大域的な特徴を捉えるための新しい相関分析手法を提案する。

まず、二つの時系列間の大域的な相関の強さの指標としてライン相関係数を定義する。また、時系列束におけるあらゆる二つの時系列の組み合わせに対するライン相関係数の算術平均を時系列束の結合度と定義する。結合度の逆数は時系列束の中に含まれる等価的に無相関な時系列の数と考えることができることから、結合度の逆数を時系列束の自由度と定義する。また、この自由度が吉川らの提案する確率過程の自由度 [1][2][3] の自然な拡張になっていることも示す。最後に、実際の気象データからこれらの特徴量を推定し、本手法の有効性を検証する。

2 時系列間のライン相関係数

N 個の時系列 $X_1(t), X_2(t), \dots, X_N(t)$ を考え、これをサイズ N の時系列束と呼ぶ。ただし、これらの時系列はそれぞれ定常かつエルゴード的であり、かつ各時系列の統計的性質は同じであるとする。このとき i 番目の時系列の時間区間 T 内における n 次モーメントの推定値 $Z_i^{(n)}(T)$ は

$$Z_i^{(n)}(T) = \frac{1}{T} \int_{-T/2}^{T/2} X_i^n(t) dt \quad (1)$$

となる。ここで $Z_i^{(n)}(T)$ と $Z_j^{(n)}(T)$ の相関係数 $\gamma_{ij}^{(n)}(T)$ は

$$\gamma_{ij}^{(n)}(T) = \frac{\text{Cov}[Z_i^{(n)}(T), Z_j^{(n)}(T)]}{\text{Var}[Z_i^{(n)}(T)]} \quad (2)$$

と表される。式 (2) において、 $T \rightarrow \infty$ の極限として $\gamma_{ij}^{(n)}$ を考える。

$$\gamma_{ij}^{(n)} = \lim_{T \rightarrow \infty} \gamma_{ij}^{(n)}(T) = \frac{\int_{-\infty}^{\infty} r_{ij}^{(n)}(\tau) d\tau}{\int_{-\infty}^{\infty} r^{(n)}(\tau) d\tau} \quad (3)$$

この $\gamma_{ij}^{(n)}$ を i 番目の時系列 $X_i(t)$ と j 番目の時系列 $X_j(t)$ の n 次ライン相関係数定義する。 n 次ライン相関係数は、 n 次モーメントを推定するという立場からみたときの二つの時系列間の全体としての相関の強さを表すものである。

時空スカラー確率場の各点で観測されるスカラー量は、それ自身がそれぞれ時系列であるから、時間的にも空間

的にも一様かつエルゴード的な時空確率場はサイズ無限の時系列束と考えることができる。この場合、これらの時系列のライン相関係数は距離の関数となり、確率場の特徴量としての相関面積や相関距離 [3] と結び付く。

3 時系列束の結合度と自由度

時系列束において重複を許すあらゆる二つの時系列の組み合わせに対するライン相関係数の算術平均 $\bar{\gamma}^{(n)}$ を

$$\bar{\gamma}^{(n)} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \gamma_{ij}^{(n)} \quad (4)$$

とおくと、 $\bar{\gamma}^{(n)}$ は n 次モーメントを推定するという立場からみたときの全時系列間の相関の強さ、即ち、この時系列束のまとまりの強さを表す。そこでこの $\bar{\gamma}^{(n)}$ を時系列束の n 次結合度と定義する。時系列束の結合度のとり得る値は $0 \leq \bar{\gamma}^{(n)} \leq 1$ であり、時系列が互いに1の相関をもつとき最大値1をとり、時系列が互いに-1の相関をもつとき最小値0をとる。また、時系列すべてが無相関であるときは $\frac{1}{N}$ となる。

ここで、結合度の逆数は時系列束の自由度と考えることができるので、時系列束の n 次自由度 $N_L^{(n)}$ を

$$N_L^{(n)} = \frac{1}{\bar{\gamma}^{(n)}} \quad (5)$$

と定義する。この $N_L^{(n)}$ は時系列束内の等価的に無相関な時系列の数を表す。この時系列束の自由度は、確率過程に対して定義される自由度 [1] の概念を多次元時系列に拡張することによって得られるものと一致する。以下にこのことを説明する。

一つの時系列に対する自由度の考え方を拡張すると、時間区間 T 内の n 次自由度 $N_G^{(n)}$ は

$$N_G^{(n)} = \frac{\text{Var}[X_i^n(t)]}{\text{Var}[\frac{1}{NT} \sum_{i=1}^N \int_T X_i^n(t) dt]} \quad (6)$$

と表せる。これは n 次モーメントに関して時系列束が時間区間 T 内に持つ等価的に無相関な点の数を表す。一方、一つの時系列が時間区間 T 内にもつ等価的に無相関な点の数 $N_T^{(n)}$ は

$$N_T^{(n)} = \frac{\text{Var}[X_i^n(t)]}{\text{Var}[\frac{1}{T} \int_T X_i^n(t) dt]} \quad (7)$$

である [1]。従って、 $N_G^{(n)}$ と $N_T^{(n)}$ の比

$$N_L^{(n)}(T) = \frac{N_G^{(n)}}{N_T^{(n)}} \quad (8)$$

は時間区間 T の N 本の時系列束の中で、等価的に無相関と見なせる時系列の本数を表すことになる。ここで、時間区間 T を無限大にしたときの $N_L^{(n)}(T)$ の極限值 $N_L^{(n)}$ は、式 (5) の n 次自由度と一致する。

4 気象データへの応用例

複数の地点で観測された気温データの解析に応用した結果を示す。対象とする時系列束は、図1に示す東京、千葉、神奈川、埼玉、愛知、大阪の54地点において1991

† 早稲田大学理工学部

School of Science and Engineering, Waseda Univ.

† 近畿大学生物理工学部

Faculty of Biology-Oriented Science and Technology, Kinki Univ.

年に観測されたアメダスの気温データである。実際の解析では、年周期、日周期を除去したデータ群を定常な時系列束とみなして解析を行なった。

図2は、全ての二つの観測地点の組み合わせに対し、1次、2次ライン相関係数を求め、観測地点間の相対的な位置との関係を示したものである。

図3は、観測地点間のライン相関係数が、東西方向距離 ξ 、南北方向距離 η の関数として、2次の実対称行列 Γ を用いて

$$r(\xi, \eta) = \exp \left\{ -\frac{1}{2} \begin{pmatrix} \xi & \eta \end{pmatrix} \Gamma \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right\} \quad (9)$$

のように表されるものとしたときの、ライン相関係数の推定値の等高線表示である。ただし、観測が行なわれた場は一様な正規確率場であると仮定し、近似には最小二乗法を用いた。

ライン相関係数より推定されるこの確率場の相関のある領域に関する特徴量[3]は表1のようになる。表1より、東西方向および南北方向の1次相関距離はそれぞれ694[km]、380[km]であり、2次相関距離はそれぞれ495[km]、266[km]であることが分かる。つまり、地上気温は南北方向より東西方向に相関が強いと言える。これらの特徴量間の関係は図4に示すとおりである。

表2はこの時系列束の1次および2次結合度、1次および2次自由度である。つまり、観測された54本の時系列の中には、平均を推定するという立場に立ったときには1.18本、分散を推定するという立場に立ったときには1.39本の等価的に無相関な時系列が存在するということになる。

5 まとめ

時系列間の相関の強さを示すn次ライン相関係数と、その時系列束のまとまりの強さを表わす結合度を定義し、更に、時系列束の持つ等価的に無相関な時系列の本数を表す時系列束の自由度を定義した。そしてこの自由度が、確率過程において定義された自由度の概念を多次元時系列に拡張することによって得られるものと同じであることを示した。また、この理論を気温データ解析に応用し、地上気温に関する相関距離などの特徴量を示した。さらに、54地点で観測された気温データからなる時系列束から、時系列束のまとまりの強さと等価的に無相関な時系列の数を推定した。

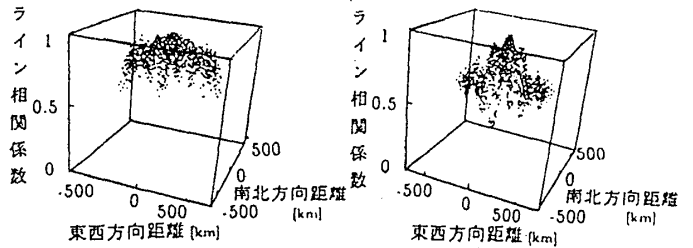
本報告で提案した相関分析法は従来の点と点との間の相関だけを用いた分析法とは異なり、複数の時系列の性質を大域的に把握するための新しい解析法であり、動画像の空間的な相関を評価するなど応用範囲も広いものと考えられる。

参考文献

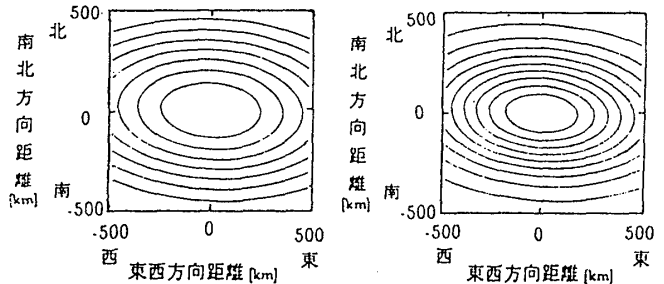
- [1] Kikkawa, S. and Ishida, M., "Number of degrees of freedom, correlation times, and equivalent bandwidths of a random process", IEEE Trans. Inform. Theory, Vol.34, No.1, pp.151-155, 1988.
- [2] 吉川 昭, "確率過程の自由度・フィッシャーの情報量・周波数-時間積", 信学論 A, Vol.J76, No.8, pp.1118-1126, 1993.
- [3] 吉川 昭, 坂本 雅人, 小林 直也, "2次元確率場の自由度と相関面積・相関距離", 信学論 A, Vol.J77, No.4, pp.711-719, 1994.



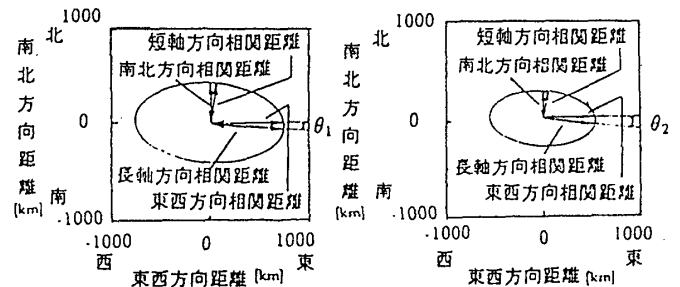
図1 観測点の分布



(a) 1次ライン相関係数 (b) 2次ライン相関係数
図2 ライン相関係数と観測地点の相対的位置の関係



(a) 1次ライン相関係数 (b) 2次ライン相関係数
図3 推定されるライン相関係数と観測地点の相対的位置の関係



(a) 1次ライン相関係数 (b) 2次ライン相関係数
図4 ライン相関係数より推定される相関距離

表1 ライン相関係数より推定されるパラメータ

	1次	2次
東西方向相関距離 [km]	694	495
南北方向相関距離 [km]	380	266
長軸方向相関距離 [km]	697	498
短軸方向相関距離 [km]	379	265
θ [rad]	-0.0619	-0.0685
相関面積 [km ²]	896000	529000

表2 時系列束の結合度、自由度

	1次	2次
時系列束の結合度	0.842	0.717
時系列束の自由度	1.18	1.39