

高次元時系列の球面性検定

京都大学大学院・情報学研究科 劉 言
早稲田大学・基幹理工学研究科 田村 百合絵
早稲田大学・理工学研究所 谷口 正信

1. はじめに

多変量標本の共分散行列の構造に関する統計理論の歴史が長い。その中で共分散行列が単位行列のスカラ一倍であるという仮説に対する検定は球面性検定という。高次元では尤度比統計量が退化するため、通常 John [1] や Nagao [3] で推奨された検定統計量が用いられる。本研究では、(i) 多変量標本が相関構造をもつ; (ii) 次元 p が標本数 n に対し、 $n \rightarrow \infty, p \rightarrow \infty$ のとき、 $p/n \rightarrow c \in (0, \infty)$ が成り立つという二条件の満たす新しい枠組みで、高次元時系列の球面性検定について調べた [2].

2. 高次元時系列の球面性検定

平均 $\boldsymbol{\mu}$ で $p \times p$ 共分散行列関数 $\mathbf{R}(t) = (R_{ij}(t); i, j = 1, \dots, p)$, スペクトル $\mathbf{f}(\lambda)$ をもつ高次元定常正規時系列 $\{\mathbf{X}(t) = (X_1(t), \dots, X_p(t))'; t \in \mathbb{Z}\}$ に対し、自己共分散行列 $\mathbf{R}(0)$ の球面性検定を考える。即ち、

$$H : \mathbf{R}(0) = \sigma^2 \mathbf{I}_p \quad \text{v.s.} \quad K : \mathbf{R}(0) \neq \sigma^2 \mathbf{I}_p$$

を考える。但し、 \mathbf{I}_p は p 次元単位行列で、 $\sigma^2 > 0$ とする。観測数 N に対し、標本共分散行列 \mathbf{S} を

$$\mathbf{S} = \frac{1}{n} \sum_{t=1}^N (\mathbf{X}(t) - \bar{\mathbf{X}})(\mathbf{X}(t) - \bar{\mathbf{X}})'$$

とする。ここでは、 $n = N - 1$ である。これを用いて、検定統計量 U を

$$U = \frac{1}{p} \text{tr} \left(\frac{\mathbf{S}}{(1/p) \text{tr} \mathbf{S}} - \mathbf{I}_p \right)^2$$

と定義する。帰無仮説 H の下で、検定統計量 U の振舞いを調べた。

通常、検定統計量 U に対し、 Δ だけ修正した $U - \Delta$ は $O(\sqrt{np})$ である。多変量標本に自己相関がない場合、 Δ は n と p のみに依存する量である。しかし、本研究 (i) の下で Δ が $\mathbf{f}(\lambda)$ にも依存する結果となった。これに加えて、 $U - \Delta$ の分散も $\mathbf{f}(\lambda)$ に依存する。これまでデータの球面性問題に対する統計的解析は往々にして $\mathbf{f}(\lambda)$ を考えず検定統計量 U が用いられてきたが、自己相関をもつ観測についても同様な操作を行うことが明らかに不適切である。

本発表はこれらの結果に加えて、共分散行列に対立仮説の構造を導入したときの数値シミュレーション結果を確認する。また、実データ解析の一例として、ニューヨーク株式市場の株式によって構成されるポートフォリオに対する球面性検定の諸結果も紹介する予定である。

References

- [1] JOHN, S. (1971). Some optimal multivariate tests. *Biometrika* **58** 123–127.
- [2] LIU, Y., TAMURA, Y. and TANIGUCHI, M. (2018). Asymptotic theory of test statistic for sphericity of high-dimensional time series. *Journal of Time Series Analysis* **39** 402–416. DOI: [10.1111/jtsa.12288](https://doi.org/10.1111/jtsa.12288).
- [3] NAGAO, H. (1973). On some test criteria for covariance matrix. *The Annals of Statistics* 700–709.