

ランダム行列におけるモーメントのゆらぎに基づく 時系列パラメータの検定

長谷川 彩子

お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻

本研究は、お茶の水女子大学の吉田裕亮氏、愛知教育大学の佐久間紀佳氏との共同研究の一部である。

1 はじめに

計算機パワーの増大に伴い、近年、ランダム行列理論は様々な分野で応用がなされている。従来のランダム行列における研究では、独立な確率変数を要素にもつ行列について、行列のサイズを無限大にした際の固有値経験分布の漸近的挙動の統計的性質が調べられていた。最近では、要素間の相関が時系列モデルで与えられるようなランダム行列に関する研究も始まっている。

本研究では、1次元および2次元 MA モデルで要素間の相関が与えられるような Wishart 型ランダム行列の固有値経験分布の極限分布のモーメントが、その時系列モデルの時系列スペクトルによって完全に記述されることを述べる。さらに、最近の Wishart 型ランダム行列のモーメントのゆらぎの厳密計算と併せて、応用の一つとして、ゆらぎの漸近正規性を用いた時系列パラメータの検定が可能であることをみる。

2 ランダム行列と自由確率論

2.1 ランダム行列

ランダム行列とは確率変数を要素にもつ行列であり、各要素が独立に標準正規分布に従う変数をもつ $N \times M$ ランダム行列を G とすると、

$$S = \frac{1}{N} G^t G \quad (1)$$

で与えられる $N \times N$ 対称ランダム行列 S を Wishart 行列という。ここで、 M, N が漸近的に $M/N \rightarrow \lambda$, $M, N \rightarrow \infty$ となるような極限をとると、Wishart 行列 S の固有値経験分布は Marchenko-Pastur 則に収束することが知られている。

Statistical testing for time series models based on the fluctuation moments of the random matrices

Ayako HASEGAWA

Division of Advanced Sciences, Graduate School of Ochanomizu University

hasegawa.ayako@is.ocha.ac.jp

2.2 自由確率論

自由確率論とは自由独立性に基づく非可換確率論のことで、自由加法的・乗法的合成積により、ランダム行列の和・積の固有値経験分布の極限分布の計算が可能となる。

3 ランダム行列のモーメントのゆらぎ

$N \times N$ ランダム行列 X_N について、 $\text{tr}(X_N^k)$ を確率変数とみる。固有値経験分布の極限分布の k 次モーメントは

$$\alpha_k = \lim_{N \rightarrow \infty} E[\text{tr}(X_N^k)] \quad (2)$$

であり、 $\text{tr}(X_N^k - \alpha_k \cdot I)$ は、ゆらぎとして捉えられる。複合 Wishart ランダム行列の場合、これらはオーダー $1/N$ の漸近正規性をもつ。すなわち、

$$N(\text{tr}(X_N^k) - \alpha_k \cdot I) = \text{Tr}(X_N^k - \alpha_k \cdot I) \quad (3)$$

は、 $N \rightarrow \infty$ のとき平均 0 の Gauss 分布に分布収束する。したがって、ゆらぎの情報は、Gauss 確率変数の族 $\{\text{Tr}(X_N^k)\}_{k \geq 1}$ であり、共分散

$$\begin{aligned} \alpha_{p,q} &= \lim_{N \rightarrow \infty} \text{Cov}[\text{Tr}(X_N^p), \text{Tr}(X_N^q)] \\ &= \lim_{N \rightarrow \infty} E[\text{Tr}(X_N^p - \alpha_p \cdot I) \cdot \text{Tr}(X_N^q - \alpha_q \cdot I)] \end{aligned} \quad (4)$$

で与えられる。

4 時系列モデルで要素間の相関が与えられるランダム行列と時系列スペクトルの関係

1次元(単変量)、または、2次元(平面)MA モデル

$$\begin{aligned} Y_t &= \sum_{i=0}^L b_i Z_{t-i}, \text{ または,} \\ Y_{i,j} &= \sum_{s=0}^V \sum_{t=0}^W c_s d_t Z_{i-s,j-t}, \quad (Z_t, Z_{i,j} \sim N(0,1) \text{ i.i.d.}) \end{aligned} \quad (5)$$

で表される時系列モデルについて、 $N \times M$ ランダム行列

$$\mathbf{X} = \begin{bmatrix} Y_1 & \cdots & Y_M \\ \vdots & \ddots & \vdots \\ Y_{(N-1)M+1} & \cdots & Y_{NM} \end{bmatrix},$$

$$\text{または } \mathbf{X} = \begin{bmatrix} Y_{1,1} & \cdots & Y_{1,M} \\ \vdots & \ddots & \vdots \\ Y_{N,1} & \cdots & Y_{N,M} \end{bmatrix} \quad (6)$$

から,

$$\widehat{\mathbf{X}} = \frac{1}{N} \mathbf{X}^t \mathbf{X} \quad (7)$$

を構成し, $\widehat{\mathbf{X}}$ の漸近比 $\lambda = M/N$ の固有値経験分布の極限分布の k 次自由キュムラントを r_k とする.

ここで, 1次元 MA モデルの場合, 自己共分散関数

$$\gamma(h) = \sum_{t=0}^{\infty} b_t b_{t+|h|} \quad (8)$$

から, Toeplitz 行列 $\mathbf{T} = (\gamma(i-j))_{ij}$ を与える. 行列 $\mathbf{T}$ の固有値経験分布の極限分布 ρ の k 次モーメント m_k は, 自己共分散関数 $\gamma(h)$ の Fourier 変換

$$f(\omega) = \sum_{h \in \mathbb{Z}} \gamma(h) e^{-\sqrt{-1}h\omega}, \quad (9)$$

すなわち, 時系列スペクトル関数を用いて,

$$m_k = \frac{1}{2\pi} \int_0^{2\pi} (f(\omega))^k d\omega \quad (10)$$

で表される. 2次元 MA モデルの場合には, 行方向, 列方向の時系列に対応する時系列スペクトルから定まる Toeplitz 行列をそれぞれ $\mathbf{T}_1, \mathbf{T}_2$, その固有値経験分布の極限分布をそれぞれ ρ_1, ρ_2 としたときの, ρ_1 と ρ_2 の自由乗法的合成積の k 次モーメント m_k を求める.

このとき

$$r_k = \lambda m_k \quad (11)$$

の関係が成立する.

5 時系列パラメータの検定

前節の関係と, 3節のモーメントのゆらぎの厳密計算を併せて, 応用の一つとして, ゆらぎの漸近正規性を用いた時系列パラメータの検定が可能であることをみる.

時系列データ $\{y_t\}_{t=1}^L$ が与えられたとき, それを MA Gauss 過程の標本路とみなす. 与えられた時系列データから, 式 (6) で表される $N \times M$ データ行列 $\mathbf{x} = (y_{(i-1)M+j})_{ij}$ を構成し, $\mathbf{x}$ の標本共分散行列 $(1/N) \mathbf{x}^t \mathbf{x}$ の 1 次および 2 次モーメント

$$m_k = \text{Tr} \left[\left(\frac{1}{N} \mathbf{x}^t \mathbf{x} \right) \right], \quad k = 1, 2 \quad (12)$$

を求め, これらを検定統計量 Z の観測値として用いる.

数値実験例

MA モデル $Y_t = 1.0Z_t + 1.0Z_{t-1} + 1.0Z_{t-2}$ に従う長さ 3600 の数値データを作る.

次に, 数値データから $N, M = 60$ のデータ行列 $\mathbf{x}$ を構成する. このとき, $\mathbf{x}$ の標本共分散行列の 2 次モーメントは $\tilde{\mu}_2 = 25.6832$ であった.

生成されたデータが設定したパラメータに従って生成されたという仮説の下, パラメータの検定を行う. このとき, モーメントのゆらぎの厳密計算から, 2 次モーメントは

$$N \left(\mu_2, \frac{\alpha_{2,2}}{N^2} \right) = N (28.0, (2.2630)^2) \quad (13)$$

に従っている.

よって, 2 次モーメントについて, Z -値は

$$z = \frac{\tilde{\mu}_2 - \mu_2}{\sqrt{\alpha_{2,2}/N^2}} = \frac{25.6832 - 28.0}{2.2630} = -1.0238 \quad (14)$$

となり, 有意水準 5% でモデルは支持される.

ここで, データ生成に用いたものとは異なるパラメータで同様の検定を行うと, 両側 5% の棄却域に入った.

上記の統計的検定は, 期待されるモーメントと分散が時系列スペクトルのみで計算されることから, AR および ARMA モデルについても同様に適用可能である.

6 おわりに

時系列モデルで要素間の相関が与えられるランダム行列と, 時系列スペクトルの関係について述べ, ランダム行列のモーメントのゆらぎの厳密計算と併せて, 時系列パラメータの検定を行った. 2次元 (平面) MA モデルパラメータの検定は, 画像のテキスト解析などに有効であると考えられる. 今後はその検定力についても議論したい.

参考文献

- [1] A. Hasegawa, N. Sakuma and H. Yoshida, *Random Matrices by MA Models and Compound Free Poisson Laws*, Probab. Math. Statist, 33(2) (2013), pp.243-254.
- [2] J.Mingo and A.Nica, *Annular noncrossing permutations and partitions, and second-order asymptotics for random matrices*, Int. Math. Res. Not, 28(2004), 1413-1460.