ランダム行列のモーメントのゆらぎを用いた時系列パラメータの検定

理学専攻・情報科学コース　長谷川彩子

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．ランダム行列理論における研究では, 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とき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る．従来では, 独立な確率変数を要素にもつランダム行列について調べられていたが, 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．

本論文では, 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 $\boldsymbol{G}$ とすると,

$$\boldsymbol{S} \ = \ \frac{1}{N}\,\boldsymbol{G}\,{}^t\boldsymbol{G} \tag{1}$$

で与えられる $N \times N$ 対称ランダム行列 $\boldsymbol{S}$ を Wishart 行列という．ここで, M, N が漸近的に $M/N \to \lambda$, $M, N \to \infty$ となるような極限をとると, Wishart 行列 $\boldsymbol{S}$ 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．

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．

本研究で用いる自由キュムラントという特性値は, 古典的な確率論におけるキュムラントの自由確率版である．自由モーメント・キュムラント公式により, k 次モーメントと k 次自由キュムラントは, 各々, k 次以下の他方の値から求められる．

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 $\boldsymbol{X}_N$ について, $\mathrm{tr}(\boldsymbol{X}_N^k)$ 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

$$\alpha_k = \lim_{N\to\infty} \mathrm{E}\big[\mathrm{tr}\big(\boldsymbol{X}_N^k\big)\big] \tag{2}$$

であり, $\mathrm{tr}\big(\boldsymbol{X}_N^k - \alpha_k \cdot \boldsymbol{I}\big)$ 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,

$$N\Big(\mathrm{tr}\big(\boldsymbol{X}_N^k\big) - \alpha_k \cdot \boldsymbol{I}\Big) \ = \ \mathrm{Tr}\big(\boldsymbol{X}_N^k - \alpha_k \cdot \boldsymbol{I}\big) \tag{3}$$

は, $N \to \infty$ 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 $\big\{\mathrm{Tr}\big(\boldsymbol{X}_N^k\big)\big\}_{k\geq 1}$ であり, 共分散

$$\begin{aligned}\alpha_{p,q} &= \lim_{N\to\infty} \mathrm{Cov}\Big[\mathrm{Tr}\big(\boldsymbol{X}_N^p\big), \mathrm{Tr}\big(\boldsymbol{X}_N^q\big)\Big] \\ &= \lim_{N\to\infty} \mathrm{E}\Big[\mathrm{Tr}\big(\boldsymbol{X}_N^p - \alpha_p \cdot \boldsymbol{I}\big) \cdot \mathrm{Tr}\big(\boldsymbol{X}_N^q - \alpha_q \cdot \boldsymbol{I}\big)\Big]\end{aligned} \tag{4}$$

で与えられる． X_N が複合実 Wishart 行列の場合, 共分散 $\alpha_{p,q}$ の, 図 1 に表す 非交叉円環型置換を用いた表現も見つけられている．

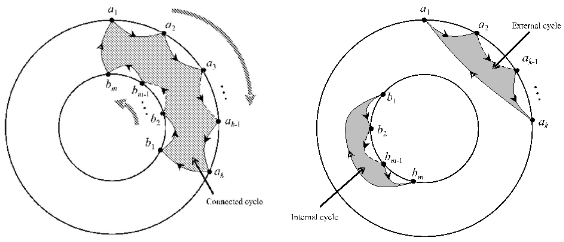

図 1. 非交叉円環型置換

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

1 次元 (単変量), または, 2 次元 (平面)MA モデル

$$Y_t = \sum_{i=0}^{L} b_i Z_{t-i}, \ \text{または,}$$

$$Y_{i,j} = \sum_{s=0}^{U}\sum_{t=0}^{V} c_s d_t Z_{i-s,j-t}, \ \ \Big(Z_t, Z_{i,j} \sim N(0,1)\,\text{i.i.d.}\Big) \tag{5}$$

で表される時系列モデルについて, $N \times M$ ランダム行列

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

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

から,

$$\boldsymbol{W} \ = \ \frac{1}{N}\,\boldsymbol{X}\,{}^t\boldsymbol{X} \tag{7}$$

を構成し, $\boldsymbol{W}$ 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

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

から, Toeplitz 行列 $\boldsymbol{T} = \big(\gamma(i-j)\big)_{ij}$ を与える．行列 $\boldsymbol{T}$ 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

$$f(\omega) \;=\; \sum_{h\in\mathbb{Z}} \gamma(h) e^{-2\pi i h\omega}, \tag{9}$$

すなわち, 時系列スペクトル関数を用いて,

$$m_k \;=\; \int_{-\frac{1}{2}}^{\frac{1}{2}} \big(f(\omega)\big)^k\, d\omega \tag{10}$$

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 $\boldsymbol{T}_1, \boldsymbol{T}_2$, 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

$$r_k \;=\; \lambda\, m_k \tag{11}$$

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統計的仮説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, $N \times M$ 行列

$$\boldsymbol{x} = \begin{bmatrix} y_1 & \cdots & y_M \\ \vdots & \ddots & \vdots \\ y_{(N-1)M+1} & \cdots & y_{NM} \end{bmatrix} \tag{12}$$

を構成し, $\boldsymbol{x}$ の標本共分散行列 $\boldsymbol{w} = (1/N)\ \boldsymbol{x}\ {}^t\boldsymbol{x}$ のモーメント

$$m_k \;=\; \frac{1}{N}\,\mathrm{Tr}\big(\boldsymbol{w}^k\big),\ \ k = 1, 2, \cdots \tag{13}$$

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．

MA パラメータは $b_0 = b_1 = b_2 = 1.0,\ b_k = 0.0\ (k \geq 3)$ であるので, 自己共分散関数は

$$\gamma(0) = 3,\ \gamma(1) = 2,\ \gamma(2) = 1,\ \gamma(k) = 0\,(k \geq 3) \tag{14}$$

となり, その Fourier 変換は

$$f(\omega) \;=\; 3 + 2(2\cos 2\pi\omega + \cos 4\pi\omega) \tag{15}$$

と考えられる．よって, 極限分布 ρ のモーメントは $m_1(\rho) = 3, m_2(\rho) = 19, m_3(\rho) = 141, \cdots$ となる．

次に, 数値データから $N, M = 60$ のデータ行列 $\boldsymbol{x}$ を構成する．このとき, $\boldsymbol{w}$ の 2 次モーメントは $\widetilde{\mu}_2 = 25.6832$ であった．

生成されたデータが設定したパラメータにしたが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．モーメントのゆらぎの厳密計算から, 2 次モーメントの期待される平均と分散は, $\mu_2 = 28,\ \alpha_{2,2}/N^2 = 18436/60^2$ によって与えられる．よって, 2 次モーメントは

$$N\Big(\mu_2,\ \frac{\alpha_{2,2}}{N^2}\Big) = N\big(28.0,\ (2.2630)^2\big) \tag{16}$$

に従っている．

よって, 2 次モーメントについて, Z-値は

$$z \;=\; \frac{\widetilde{\mu}_2 - \mu_2}{\sqrt{\alpha_{2,2}/N^2}} \;=\; \frac{25.6832 - 28.0}{2.2630} \;=\; -1.0238 \tag{17}$$

となり, 有意水準 5% の採択域 $|Z| < 1.96$ に入るため, 仮説は支持される．

ここで, データ生成に用いたものとは異なるパラメータ, 例えば, $b_0 = b_1 = 1.0,\ b_2 = 0.5,\ b_k = 0.0\ (k \geq 3)$ や, $b_0 = b_1 = b_2 = b_3 = 1.0,\ b_k = 0.0\ (k \geq 4)$ で同様の検定を行ったところ, それぞれ両側 5% の棄却域に入った．

上記の検定は, 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．生成されたデータが設定したパラメータにしたがって生成されたという仮説の下, データ生成に用いたパラメータの検定を行った場合には, 検定統計量の観測値は採択域に入り, データ生成に用いたパラメータとは異なるパラメータの検定を行った場合には棄却域に入った．2 次元 (平面)MA モデルパラメータの検定は, 画像のテクスチャ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.