

Wigner 型ランダム行列による モーメントのゆらぎのデータ解析への応用

須田羽奈子^{†1,a)}

概要：近年，データ蓄積量が膨大になっており，それに伴い少ない計算量でデータ解析の精度をあげることが求められている．また，近年ではランダム行列理論を用いたデータ解析への応用も進められている．本研究では，時系列データを Wigner 型行列に配置したものを相関のあるランダム行列とみなし，このゆらぎを用いた，時系列モデルのパラメータ検定手法を提案する．

キーワード：ランダム行列，Wigner 行列，時系列モデル

1. はじめに

近年，データの蓄積量が膨大になっており，それに伴い少ない計算量でデータ解析の精度をあげることが求められている．従来のランダム行列に関する研究では，ランダム行列の固有値の経験分布が調べられており，最近では，時系列データ解析への応用も研究されている．例えば，Wishert 型のランダム行列の場合，各要素が独立でない場合の固有値の経験分布やそのゆらぎも調べられている．しかし，Wigner 型におけるゆらぎはまだ正確には知られていない．Wigner 型では，Wishert 型の約半分のデータ量でデータ解析が出来るため，Wigner 型での固有値のゆらぎを知ることは重要である．

本研究では，各要素が独立でない Wigner 型ランダム行列の固有値経験分布のモーメントとそのゆらぎを表す同時分布の共分散行列を推定し，統計的検定への応用を試みる．

2. ランダム行列

2.1 ランダム行列とは

ランダム行列とは行列要素が確率変数として与えられている行列のことである．各要素が確率変数であるため，固有値も確率変数である．また，ランダム行列の固有値の漸近的性質に関する理論はランダム行列理論 (RMT) と呼ばれている．代表的なものに Wigner 半円則や Marchenko-Pastur 則がある．

2.2 Wigner 半円則

実対称ランダム行列の固有値経験分布は行列のサイズ $n \rightarrow \infty$ で密度関数

$$\frac{\sqrt{4\sigma^2 - x^2}}{2\pi\sigma^2}$$

に収束することは古くから知られている．(σ^2 は分散)

2.3 固有値のモーメント

行列の各要素が独立な Wigner 型ランダム行列の場合， $n \times n$ のランダム行列 A の k 次モーメントは $\text{tr}(A^k)$ ，すなわち A の固有値の k 乗の平均であり， k 次モーメントは $n \rightarrow \infty$ で，以下の値に収束する．

$$m_k = \begin{cases} 0 & \text{if } k = 2n + 1, \\ \frac{(2n)!}{(n+1)!n!} & \text{if } k = 2n. \end{cases}$$

($\frac{(2n)!}{(n+1)!n!}$ はカタラン数である)

本研究では，Wigner 型において各要素間にある相関を持つ場合のモーメントにまず着目する．

3. モーメント値とゆらぎの推定

3.1 理論値の導出

本研究では各要素が独立でない場合のランダム行列として，要素間の相関が時系列 MA モデルで与えられる場合について考える．MA モデルのパラメータを c_i とし

$$Y_i = \sum_{k=0}^{\infty} c_k Z_k, \quad (\text{ただし, } Z_k \sim N(0, 1) \text{ i.i.d.})$$

に従う MA 過程を生成し，これを Wigner 型に並べた

$$X = \begin{pmatrix} Y_1 & Y_2 & \dots & Y_n \\ Y_2 & Y_{n+1} & \dots & Y_{2n-1} \\ \vdots & \vdots & \ddots & \vdots \\ Y_n & Y_{2n-1} & \dots & Y_{n(n+1)/2} \end{pmatrix}$$

^{†1} 現在，お茶の水女子大学院
Presently with Ocha University

^{a)} suda.hanako@is.ocha.ac.jp

をランダム行列とみなす。

このランダム行列 X の固有値経験分布の極限分布 μ のモーメントは、以下のように求められる。

まず X は Wigner 型なので μ の奇数次モーメントは 0 となる。 μ の偶数次のモーメントに関しては、次のようにして求められる。

時系列の離散自己相関関数 γ を用いて Toeplitz 行列 Γ_c を構成する。すなわち、

$$\gamma_c(h) = \sum_{j=0}^{\infty} c_j c_{j+|h|}, \quad \Gamma_c = (\gamma_c(i-j))_{i,j}.$$

行列 Γ_c の固有値経験分布の極限分布 ρ とし、自由乗法的合成積を $\nu = \rho \boxtimes \rho$ とおく。

このとき、ランダム行列 X の固有値経験分布の極限分布 μ の $2k$ 次自由キウムラント α_{2k} が ν の k 次モーメントの平方根で与えられることになる。もちろん、 μ の奇数次キウムラントは 0 である。 X の偶数次モーメントは 自由モーメント・キウムラント公式を用いて

$$\begin{aligned} m_2 &= \alpha_2, \\ m_4 &= \alpha_2^2 + \alpha_4, \\ m_6 &= \alpha_2^3 + 3\alpha_2\alpha_4 + \alpha_6, \\ m_8 &= \alpha_2^4 + 6\alpha_2^2\alpha_4 + 2\alpha_4^2 + 2\alpha_2\alpha_6 + \alpha_8 \end{aligned}$$

と表すことができる。このようにしてモーメントを求めることは可能である。

4. ゆらぎ

一般に X のモーメントのゆらぎは、

$$\begin{aligned} \beta_{p,q} &= \lim_{n \rightarrow \infty} \text{Cov}(\text{Tr}(X_n^p), \text{Tr}(X_n^q)) \\ &= \lim_{n \rightarrow \infty} \text{E}[\text{Tr}(X_n^p - m_p \mathbf{1}) \cdot \text{Tr}(X_n^q - m_q \mathbf{1})] \end{aligned}$$

と捉えることができる。Wigner 行列のゆらぎに関する厳密解はまだ知られていないので、本研究では n が大きくなるにつれて $\beta_{p,q}$ が一般に $O(n^{-2})$ で収束することを用いてゆらぎの予測値を推定した。

4.1 ゆらぎ推定の概要

各行列のサイズ $n = 50r$ ($1 \leq r \leq 10$) においてそれぞれのランダム行列の 1 次から 8 次モーメントまでを求めて 8 変量データを得る。これを 1000 回繰り返し、8 変量の 1000 サンプルを得る。この、実験で得られた r 個の 1000×8 共分散行列の各 (p, q) について n と共分散行列の値の両対数をプロットし、予測値を求め、ゆらぎを推定する。この時 $\beta_{p,q}$ は $O(n^{-2})$ で収束するため、回帰直線の切片が $\log \beta_{p,q}$ の予測値となる。 $p, q = 8$ までの各要素について予測値を算出した (図 1)。

5. 時系列パラメータの検定

モーメントとそのゆらぎの応用として、時系列モデルの検定が挙げられる。MA 時系列モデルに基づく標本 y_t を

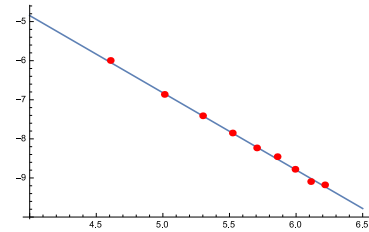


図 1 実験共分散行列 (2,2) の収束

$N \times N$ の Wigner 型に先と同様に配置する。 x の k 次モーメントについて平均は 3 節で求められ、分散は 4 節のゆらぎの、 $\beta_{2,2}$ である。ゆらぎは漸近的にガウス分布で与えられることが知られているので、 z 検定を行う。 x の標本 k 次モーメント

$$m_k = \text{Tr} \left[\frac{1}{\sqrt{n}} x^k \right]$$

を Z 検定量の観測値として有意水準 5% で検定を行った。

5.1 時系列モデル検定の数値実験

MA モデル $y = 1.0z_t + 1.0z_{t-1}$ に従うデータを生成し、 $n = 50$ の Wigner 型データ行列 x に配置する。本実験では $k = 2$ とし、2 次モーメントで実験を行う。生成した MA モデル x の標本 2 次モーメントは 8.45 となった。理論値の 2 次モーメント、数値計算から推定された分散を用いて z 値を求め、検定を行う。

検定力を確かめるために、生成されたデータが設定したパラメータに従い生成されたとした場合と、そうでない別のパラメータによって生成された場合の 2 つの検定を行なった。以下の表は 2 種類のパラメータで検定を行なった結果である。この表より正しいパラメータでの場合は z 値

パラメータ値	$c_1 = c_2 = 1$	$c_1 = c_2 = c_3 = 1$
標本 2 次モーメント	8.45	21.82
従う分布 $N(m_2, \beta_{2,2}/N^2)$	$N(8.45, (0.96)^2)$	$N(21.82, (0.96)^2)$
z 検定量 $\frac{\hat{m}_2 - m_2}{\sqrt{\beta_{2,2}/N^2}}$	-1.25	12.64

の絶対値が 1.96 より小さくなり、誤ったパラメータでの場合は z 値の絶対値が 1.96 より大きくなったため、正しい採択が出来ることが確認された。

6. まとめと今後の課題

本研究において各要素が MA モデルで相関がある場合の Wigner 型におけるランダム行列のモーメントの可能性について議論した。また、このモーメントのゆらぎを数値実験から推定し、それを用いたパラメータの検定手法を検討した。従来の研究では Wishert 型における各要素が独立でない場合のモーメント、ゆらぎが既知であったが、本研究において Wigner 型に拡張できたことで、より少ないデータ数で同等の精度のデータ解析が行える可能性が示唆された。今後は、ゆらぎの厳密解を与えることが課題としてあげられる。