メッセージ伝播法の入門から最先端まで
第5回講義資料
ランダム行列理論入門

九州大学
令和元年10月10日

豊橋技術科学大学
電気・電子情報工学系
准教授　竹内啓悟

ランダム行列理論

ランダム行列

確率変数を要素として持つ行列

目的

（対称）ランダム行列の固有値や固有ベクトルが、どんな確率分布に従うかを解明すること。

方法論

可換変数に関する古典的確率論を非可換変数に拡張した自由確率論[5-1, 2]を数学的道具とする。

自由・・・古典的確率論の独立性に対応する概念

[5-1] F. Hiai and D. Petz, *The Semicircle Law, Free Random Variables, and Entropy*, Amer. Math. Soc., 2000.

[5-2] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications,* Now Pub., 2004.

直交不変性

$\boldsymbol{X}$を任意のランダム行列とする。

左直交不変性

$\boldsymbol{\Phi X}$が定義可能で、$\boldsymbol{X}$と独立なすべての直交行列$\boldsymbol{\Phi}$に対して、以下の条件が成立するとき、$\boldsymbol{X}$は左直交不変であると言う。

$$\boldsymbol{X} \sim \boldsymbol{\Phi X}.$$（両辺の分布が同一）

右直交不変性

$\boldsymbol{X \Psi}$が定義可能で、$\boldsymbol{X}$と独立なすべての直交行列$\boldsymbol{\Psi}$に対して、以下の条件が成立するとき、$\boldsymbol{X}$は右直交不変であると言う。

$$\boldsymbol{X} \sim \boldsymbol{X \Psi}.$$

ハール直交行列

$N \times N$直交行列全体を表す空間を$\mathcal{O}_N$と書く。

ハール直交行列

左直交不変な直交行列 $\boldsymbol{U} \in \mathcal{O}_N$ は、ハール分布すると言い、
ハール直交行列と呼ばれる。

性質

ハール直交行列$\boldsymbol{U} \in \mathcal{O}_N$は、右直交不変である。

$\because$ $\boldsymbol{U}$と独立な$\boldsymbol{\Psi} \in \mathcal{O}_N$に対して、$\boldsymbol{U\Psi} \in \mathcal{O}_N$は左直交不変である。

左直交不変なハール確率測度の一意性から、

$$\boldsymbol{U\Psi} \sim \boldsymbol{U}.$$

したがって、$\boldsymbol{U}$は右直交不変である。 ■

特異値分解に基づく分布表現

$\boldsymbol{U} \in \mathcal{O}_M$、$\boldsymbol{\Sigma} \in \mathbb{R}^{M\times N}$、$\boldsymbol{V} \in \mathcal{O}_N$をそれぞれ直交行列、対角成分に非負要素を持って他の要素は0の行列、$\boldsymbol{U\Sigma}$と独立なハール直交行列とする。行列$\boldsymbol{X} \in \mathbb{R}^{M\times N}$の分布が$\boldsymbol{U\Sigma V}^{\mathrm{T}}$の分布と等しいための必要十分条件は、$\boldsymbol{X}$が右直交不変であることである。

必要性　$\boldsymbol{X} \sim \boldsymbol{U\Sigma V}^{\mathrm{T}}$を仮定する。$\boldsymbol{\Psi}$を$\boldsymbol{X}$と独立な直交行列として、

$$\boldsymbol{X\Psi} \sim \boldsymbol{U\Sigma V}^{\mathrm{T}}\boldsymbol{\Psi} = \boldsymbol{U\Sigma}\left(\boldsymbol{\Psi}^{\mathrm{T}}\boldsymbol{V}\right)^{\mathrm{T}} \sim \boldsymbol{U\Sigma V}^{\mathrm{T}} \sim \boldsymbol{X}. \qquad \blacksquare$$

十分性　$\boldsymbol{X}$の特異値分解を$\boldsymbol{X} = \boldsymbol{U\Sigma V}_0^{\mathrm{T}}$とする。$\boldsymbol{V}$を$\boldsymbol{X}$と独立なハール直交行列とすると、$\boldsymbol{V}$は$\boldsymbol{U\Sigma}$と独立である。

$\boldsymbol{X}$の右直交不変性から、

$$\boldsymbol{X} \sim \boldsymbol{XV}^{\mathrm{T}} = \boldsymbol{U\Sigma V}_0^{\mathrm{T}}\boldsymbol{V}^{\mathrm{T}} \sim \boldsymbol{U\Sigma V}^{\mathrm{T}}. \qquad \blacksquare$$

注意　特異値分解が一意の場合、$\boldsymbol{X} = \boldsymbol{U\Sigma V}^{\mathrm{T}}$を主張できる。

経験固有値分布

経験固有値分布

対称行列$\boldsymbol{X} \in \mathbb{R}^{N\times N}$の固有値を$\{\lambda_n \in \mathbb{R}\}_{n=1}^{N}$とする。

$$\rho_{\boldsymbol{X}}(\lambda) = \frac{1}{N}\sum_{n=1}^{N} 1(\lambda_n \le \lambda).$$

$\lambda \in \mathbb{R}$以下の固有値の数の割合

固有値が確率変数なので、固有値分布自体も確率的である。

経験(二乗)特異値分布

行列$\boldsymbol{X} \in \mathbb{R}^{M\times N}$に対して、$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$の固有値を$\{\lambda_n \ge 0\}_{n=1}^{N}$とする。

$$\rho_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(\lambda) = \frac{1}{N}\sum_{n=1}^{N} 1(\lambda_n \le \lambda).$$

二乗特異値の数の割合

固有値分布のモーメント系列

対称行列$\boldsymbol{X} \in \mathbb{R}^{N\times N}$の経験固有値分布$\rho_X(\lambda)$の$k$次モーメント

$$\mu_k = \int \lambda^k d\rho(\lambda) = \int \lambda^k \frac{d}{d\lambda}\frac{1}{N}\sum_{n=1}^{N} 1(\lambda_n \leq \lambda)\, d\lambda$$

$$= \frac{1}{N}\sum_{n=1}^{N}\int \lambda^k \delta(\lambda - \lambda_n) d\lambda = \frac{1}{N}\sum_{n=1}^{N}\lambda_n^k .$$

性質

$$\mu_k = \frac{1}{N}\mathrm{Tr}\left(\boldsymbol{X}^k\right).$$

$\because$　$\boldsymbol{\Lambda} = \mathrm{diag}\{\lambda_1, \dots, \lambda_N\}$として、固有分解$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}}$を使うと、

$$\frac{1}{N}\mathrm{Tr}\left(\boldsymbol{X}^k\right) = \frac{1}{N}\mathrm{Tr}\left(\boldsymbol{U}\boldsymbol{\Lambda}^k\boldsymbol{U}^{\mathrm{T}}\right) = \frac{1}{N}\mathrm{Tr}\left(\boldsymbol{\Lambda}^k\right) = \frac{1}{N}\sum_{n=1}^{N}\lambda_n^k = \mu_k.$$

ガウス行列の漸近特異値分布

$\boldsymbol{X} \in \mathbb{R}^{M\times N}$を独立な$\mathcal{N}(0, M^{-1})$に従う要素を持つガウス行列とする。

注意

ガウス行列$\boldsymbol{X}$は、両側直交不変である。

定理5.1

Wishart行列$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$の経験固有値分布$\rho_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(\lambda)$は、$\delta = M/N$を固定して$M, N \to \infty$とした大システム極限において、Marchenko・Pastur（MP）分布$\rho_{\mathrm{MP}}(\lambda)$に概収束する。

$$\frac{d\rho_{\mathrm{MP}}}{d\lambda}(\lambda) = (1-\delta)^{+}\delta(\lambda) + \frac{\delta\sqrt{(\lambda-\lambda_{-})^{+}(\lambda_{+}-\lambda)^{+}}}{2\pi\lambda}.$$

ただし、$\lambda_{\pm} = \left(1 \pm \delta^{-1/2}\right)^{2}$と定義した。

MP分布のサポートは$[\lambda_{-}, \lambda_{+}] \cup \{0\}$である。

Stieltjes変換

経験固有値分布$\rho_X(\lambda)$のStieltjes変換

$$S_X(z) = \int \frac{d\rho_X(\lambda)}{\lambda - z} = \frac{1}{N}\sum_{n=1}^{N} \frac{1}{\lambda_n - z}, \quad z \in \mathbb{C}^+ = \{z \in \mathbb{C}: \mathrm{Im}[z] > 0\}.$$

べき級数展開

$|z| > \max\{\lambda_1, \dots, \lambda_N\}$に対して、

$$S_X(z) = -\frac{1}{zN}\sum_{n=1}^{N} \frac{1}{1 - \lambda_n/z} = -\frac{1}{zN}\sum_{n=1}^{N}\sum_{k=0}^{\infty} \frac{\lambda_n^k}{z^k} = -\frac{1}{z}\sum_{k=0}^{\infty} \frac{\mu_k}{z^k}.$$

Stieltjes変換は、経験固有値分布を定義するのに必要なすべての情報を持っている。

TOYOHASHI UNIVERSITY OF TECHNOLOGY

Stieltjes変換の性質

累積分布$\rho(\lambda)$のStieltjes変換を$S(\lambda)$とする。

性質１　Stieltjes変換の虚部は正である。

$\because$
$$S(x+iy)=\int\frac{(\lambda-x)d\rho(\lambda)}{(\lambda-x)^2+y^2}+i\int\frac{yd\rho(\lambda)}{(\lambda-x)^2+y^2},\qquad y>0.$$

性質２　$\rho(\lambda)$が微分可能ならば、
$$\rho'(\lambda)=\frac{1}{\pi}\lim_{\omega\downarrow 0}\mathrm{Im}[S(\lambda+i\omega)].$$

性質３　$S(z)$の孤立特異点$z=a\in\mathbb{R}$での留数が$p<0$ならば、
$\rho(\lambda)$は点$\lambda=a$で$|p|$だけ不連続に増加する。

$\because$
$$\frac{p}{z-a}=\int\frac{-p\delta(\lambda-a)d\lambda}{\lambda-z}.$$

η変換

経験固有値分布$\rho_{\boldsymbol{X}}(\lambda)$の η 変換

$$\eta_{\boldsymbol{X}}(z) = \int \frac{d\rho_{\boldsymbol{X}}(\lambda)}{1 + z\lambda} = \frac{1}{N}\sum_{n=1}^{N} \frac{1}{1 + z\lambda_n} \quad \text{for } z \geq 0.$$

η 変換のべき級数表示

原点の近傍において、

$$\eta_{\boldsymbol{X}}(z) = \sum_{k=0}^{\infty} \mu_k (-z)^k .$$

Stieltjes変換と η 変換の関係

$$\eta_{\boldsymbol{X}}(z) = \frac{1}{z}\int \frac{d\rho_{\boldsymbol{X}}(\lambda)}{\lambda + z^{-1}} = \frac{S_{\boldsymbol{X}}(-z^{-1})}{z}.$$

R変換

経験固有値分布$\rho_{\boldsymbol{X}}(\lambda)$のR変換

$$R_{\boldsymbol{X}}(z) = S_{\boldsymbol{X}}^{-1}(-z) - \frac{1}{z}.$$

R変換とη 変換の関係

$$\eta_{\boldsymbol{X}}(z) = \frac{1}{1 + zR_{\boldsymbol{X}}(-z\eta_{\boldsymbol{X}}(z))}.$$

$\because$ $s = S_{\boldsymbol{X}}(-z^{-1})$とおくと、$R_{\boldsymbol{X}}(-s) = s^{-1} - z^{-1}$なので、

$$\frac{s}{z} = \frac{1}{1 + zR_{\boldsymbol{X}}(-s)}.$$

最後に、η 変換の定義$\eta_{\boldsymbol{X}}(z) = s/z$を代入すれば良い。

定理5.1の証明の手順

手順1　経験固有値分布$\rho_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(\lambda)$のStieltjes変換$S_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z)$を大システム極限で評価する。

- Wishart行列をランク1行列の和として表現する。

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \sum_{m=1}^{M} \boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m, \quad \boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1 \\ \vdots \\ \boldsymbol{x}_M \end{bmatrix}.$$

- ガウス行列の場合に特有のR変換の漸近加法性を使って、$R_{\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m}(z)$から$R_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z)$を計算する。

$$R_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z) - \sum_{m=1}^{M} R_{\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m}(z) \to 0.$$

$\{\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m\}$の漸近自由性

- R変換$R_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z)$からStieltjes変換を計算する。

手順2　得られた漸近的なStieltjes変換を逆変換する。

$\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m$のR変換の計算

行列$\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m$は、固有値$a=\|\boldsymbol{x}_m\|^2$と$N-1$個の重複固有値0を持つ。

$$\eta_{\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m}(\tilde{z})=\frac{1}{N}\frac{1}{1+a\tilde{z}}+\frac{N-1}{N}\quad \text{for } \tilde{z}\geq 0.$$

$z=-\tilde{z}\eta_{\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m}(\tilde{z})$とおき、$\tilde{z}\geq 0$に関して解くと、

$$\tilde{z}=\frac{-(az+1)+\sqrt{(1-az)^2+4az/N}}{2(1-1/N)a}$$

$$=-z+\frac{az^2}{N(1-az)}+\mathcal{O}(N^{-2}).$$

ここで、十分大きなNに対して$\tilde{z}\geq 0$となる解を選んだ。

これをη変換とR変換の関係式に代入すると、

$$R_{\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m}(z)=\frac{1}{\tilde{z}\eta_{\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{x}_m}(\tilde{z})}-\frac{1}{\tilde{z}}=\frac{\|\boldsymbol{x}_m\|^2}{N(1-\|\boldsymbol{x}_m\|^2 z)}+\mathcal{O}(N^{-2}).$$

$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$の各種変換の計算

R変換　ガウス行列の場合に特有のR変換の漸近加法性を使う。

$$R_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z) = \sum_{m=1}^{M} \frac{\|\boldsymbol{x}_m\|^2}{N(1-\|\boldsymbol{x}_m\|^2 z)} + o(1) \xrightarrow{\text{a.s.}} \frac{\delta}{\delta - z}.$$

Stieltjes変換　R変換の定義から、$\dfrac{\delta}{\delta+z} = S^{-1}_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z) + \dfrac{1}{z}.$

これをzについて解き、適切な符号の解を選ぶと、

$$S_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z) = \frac{1-\delta^{-1}-z-\sqrt{f(z)}}{2\delta^{-1}z}, \quad \mathrm{Im}\left[\sqrt{f(z)}\right] < 0,$$

$$f(z) = z^2 - 2z(1+\delta^{-1}) + (1-\delta^{-1})^2.$$

η 変換　R変換とη 変換の関係式に上記を代入して整理すると、

$$z\eta^2_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z) + (\delta z - z + \delta)\eta_{\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}}(z) - \delta = 0.$$

漸近固有値分布の導出

- MP分布の離散成分は、Stieltjes変換の性質3から従う。

$$\lim_{z\to 0} z S_{X^{\mathrm{T}}X}(z) = \frac{1-\delta^{-1}-\sqrt{(1-\delta^{-1})^2}}{2\delta^{-1}} = \begin{cases} \delta - 1 & \text{for } \delta \le 1 \\ 0 & \text{for } \delta > 1. \end{cases}$$

- 連続成分の評価

平方根の中身は$f(\lambda) = (\lambda - \lambda_-)(\lambda - \lambda_+)$なので、

$$\lim_{\omega \downarrow 0} \mathrm{Im}\left[\sqrt{f(\lambda + i\omega)}\right] = \begin{cases} -|f(\lambda)|^{1/2} & \text{for } \lambda \in (\lambda_-, \lambda_+), \\ 0 & \text{otherwise.} \end{cases}$$

$\mathrm{Im}\left[\sqrt{f(\lambda+i\omega)}\right] < 0$となる解を選択した。

Stieltjes変換の性質2を使うと、連続成分の表現を得る。

$$\frac{1}{\pi}\lim_{\omega\downarrow 0} \mathrm{Im}[S_{X^{\mathrm{T}}X}(\lambda + i\omega)] = \begin{cases} \dfrac{|f(\lambda)|^{1/2}}{2\pi\delta^{-1}\lambda} & \text{for } \lambda \in (\lambda_-, \lambda_+), \\ 0 & \text{otherwise.} \end{cases}$$

■