橋口研究室 理学部第一部 応用数学科

- 当研究室では，統計科学，数理統計学，多変量解析の研究をしています．現在は，特に，ランダムな行列の固有値分布に興味があります．
- 固有値の分布を求める上では，大きく 2 つの方法があります．
  1. あくまでも正確な分布を求めて，正確な分布の表現をもとに数値計算する方法，
  2. より扱いやすい分布で近似する方法．数値計算も容易．
- 固有値分布の応用としては，多変量解析（主成分分析，分散分析，球形検定など），社会学におけるネットワーク分析，株式のポートフォリオ問題，無線通信などの信号処理などがあります．

1. 多変量解析と固有値分布

多変量解析での正規性の仮定

- $\mathbf{x}_1, \ldots, \mathbf{x}_n$: 多変量正規分布 $N_m(0, \Sigma)$ からのサンプル
- $\mathbf{W} = \mathbf{X}\mathbf{X}^\top$ $(\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n))$　$\mathbf{S} = \mathbf{W}/n$: 共分散行列
- $\ell_1 > \ell_2 > \cdots > \ell_m > 0$: W の固有値 (確率的に変化)
- $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$: Σの固有値 (定数)

(1) 最大固有値 ℓ_1 の分布関数 ($\rightarrow$ 無限級数になっている)

$$\Pr[\ell_1 < x] = C \exp\left(-\frac{x}{2}\operatorname{tr}\Sigma^{-1}\right) x^{\frac{1}{2}nm} {}_1F_1\left(a; b; \frac{x}{2}\Sigma^{-1}\right)$$
$$= C \exp\left(-\frac{x}{2}\operatorname{tr}\Sigma^{-1}\right) x^{\frac{1}{2}nm} \sum_{k=0}^{\infty} \frac{x^k}{2^k} \sum_{\kappa \vdash k} \frac{(a)_\kappa}{(b)_\kappa\, k!} \mathcal{C}_\kappa(\Sigma^{-1})$$

(ただし $a = \frac{m+1}{2}, b = \frac{n+m+1}{2}$，$\mathcal{C}_\kappa(\mathbf{A})$ はゾーナル多項式)

(2) 最大固有値 ℓ_1 の正規分布での近似 (nが十分大きいとき)

$$\ell_1/n \sim N(\lambda_1, 2\sqrt{\lambda_1})$$

(3) 最大固有値 ℓ_1 のχ^2分布での近似 (nが十分大きい＋いくつかの条件)

$$\ell_1/\lambda_1 \sim \chi^2_n$$

(2), (3) の数値計算は容易だが，(1)の数値計算は大変．収束も遅い．

(1) の数値例 $n = 10$，$\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 5 \end{pmatrix}$

$$8.01667 \times 10^{-12} e^{-6x} x^{10} (1 + 0.138462x + 0.0125641x^2 + \cdots + 2.25277 \times 10^{-307} x^{149} + 8.44807341 \times 10^{-310} x^{150} + O(x^{151}))$$

- x^{150} の係数は 10^{-310} と非常に小さい数．
- 次元 m が大きくなると，x^{150} まで数値計算しても，分布関数は 1 に収束しなくなる．

実ウィシャート行列固有値分布の既存の研究結果と課題

固有値分布	最大ℓ_1	$\ell_2 \sim \ell_{m-1}$	最小ℓ_m
精密	Sugiyama (1967)	?	Khatri (1972)
正規近似	Anderson (1963), Sugiura (1973)		
χ^2近似	Takemura & Sheena (2005)		
χ^2積近似	加藤，橋口 (2013)		

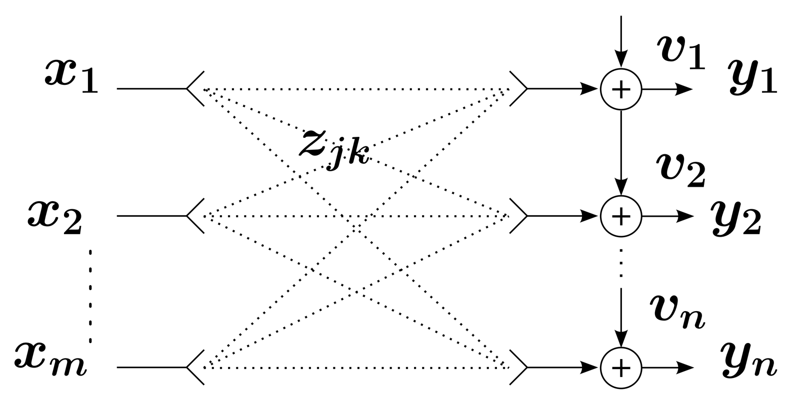
- 正確な分布の表現は，最大，最小固有値のみで得られている．間の固有値分布は研究課題
- 正確な分布の表現に基づく数値計算法の開発 (Koev & Edelman 2006, Hasiguchi and Niki 2006, Hashiguchi Numata Takayama & Takemura 2013).
- (1) の数値計算例で，無限級数を打ち切ったときの誤差評価も研究課題 (Koev & Edelman 2006)
- 新しい近似方法の開発．極限操作にバリエーションがある．
- 実際のデータ解析での利用

2. 複素数の多変量正規分布の理論と応用

- 複素数の正規分布の理論は，Wooding (1956) で信号処理での利用を目的に始まる．
- Goodman (1963) は実数の正規分布の理論との対比で，複素数の正規分布の理論を紹介．
- Telatar (1999) で無線通信での通信路容量の解析へ応用．

MIMO

- MIMO:(Multiple-input multiple-output) $\cdots$ 複数のアンテナから,異なる信号を同時に同じ周波数に重ねて送信を行う無線通信の技術.



- 入力信号 $\mathbf{x} = (x_1 \cdots x_m)^\top$，ノイズ $\mathbf{v} = (v_1 \cdots v_n)^\top$　$\mathbf{x}$，　$\mathbf{v}$ は複素正規分布にしたがう．$E\,|\mathbf{x}|^2 \leq \rho$; ρ: 電力
- 出力信号 $\mathbf{y} = (y_1 \cdots y_n)^\top$：　$\mathbf{y} = \mathbf{Z}\mathbf{x} + \mathbf{v}$

シャノン流の通信路容量 $\cdots$ 通信路で伝送される情報量

$$\text{通信路容量 } C = \mathrm{E}\left[\log\det\left(I_m + \frac{\rho}{m}\mathbf{Z}^\dagger\mathbf{Z}\right)\right]$$

$\mathbf{Z}$:ガウス通信路($\mathbf{Z}$の要素が複素正規分布に従ってランダムに変化)

- $\mathbf{W} = \mathbf{Z}^\dagger\mathbf{Z}$ 複素のウィシャート分布にしたがう
- $\ell_1 > \ell_2 > \cdots > \ell_m$: W の固有値 (確率的に変化)
- $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m > 0$: Σの固有値 (定数)

$$C = \sum_{k=1}^{m} \mathrm{E}\left[\log\left(1 + \frac{\rho}{m}\ell_k\right)\right]$$

$\Rightarrow$ 通信路容量 C を求めるためには，固有値 $\ell_1, \ldots, \ell_m$ の分布が必要

- $\ell_1, \ell_2, \cdots, \ell_m$ に順序をつけない場合，

$$C = m\,\mathrm{E}\left[\log\left(1 + \frac{\rho}{m}\ell_1\right)\right]$$

- $\lambda_1, \lambda_2, \cdots, \lambda_m$ が十分離れている＋いくつかの条件がある場合，$\ell_1, \ldots, \ell_m$ は独立であって，各 ℓ_k は形状パラメータ $n-k+1$，尺度パラメータ λ_k のガンマ分布で近似できる．(実数の場合の類似)

$$\ell_k \sim \mathrm{Gamma}(n-k+1, \lambda_k)$$

  - 形状パラメータ α，尺度パラメータ β のガンマ分布 $\mathrm{Gamma}(\alpha, \beta)$ の密度関数

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)$$

3 スティーフェル多様体の理論と方向統計学への応用

- $V_{m,n} = \{H_1\ (m \times n) \mid H_1{}^\top H_1 = I_m\}$: スティーフェル多様体 m 次元の n 本のベクトルが全て直交し，長さが 1 となっている．
- $m = 1$ のとき単位超球面 $\mathbb{S}^{m-1}$
- $\mathrm{Vol}(V_{m,n})$: $V_{m,n}$ の体積

$$\mathrm{Vol}(V_{m,n}) = \frac{2^m \pi^{mn/2}}{\Gamma_m(\frac{1}{2}n)}$$

- 固有値の同時分布は $V_{m,n}$ や $O(m) = V_{m,m}$ 上での積分を利用して得られる

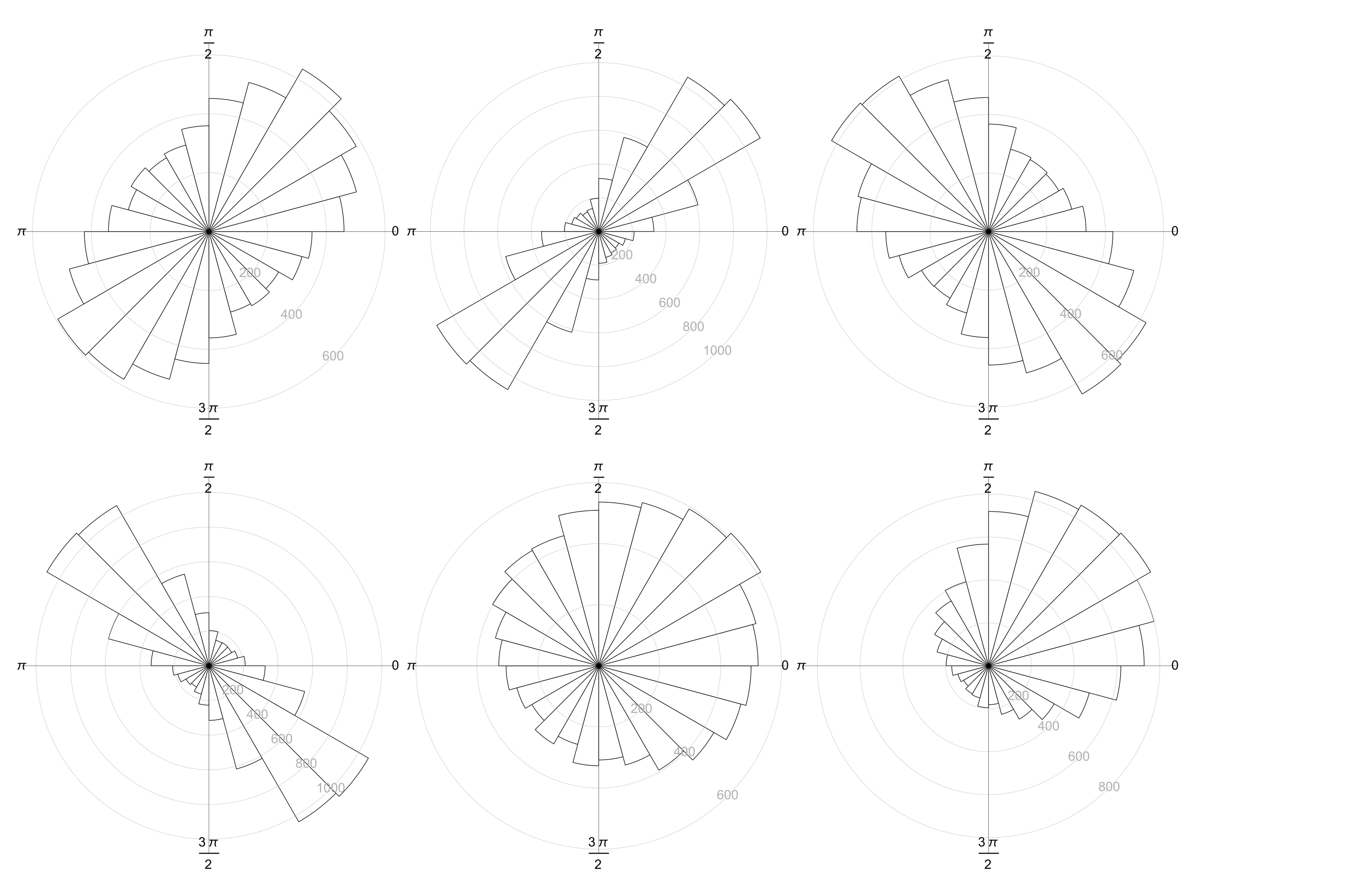
$$f(\ell_1, \ldots, \ell_m) = \int_{H \in O(m)} f_W(HLH^\top)(dH)$$

$L = \mathrm{diag}(\ell_1, \ldots \ell_m)$ は $m \times m$ ランダム行列 W の固有値の行列

超球面上の一様性検定問題

- 単位超球面$\mathbb{S}^{m-1}$上のデータ
  $\mathbb{S}^{m-1} = \{\ X \in \mathbb{R}^m : \|X\| = X'X = 1\}$.
  例)　風向，移動方向，周期性のあるデータなど．

円周上のデータ（方向データ）のバラ図による表現 ($m = 2$)



球面上の一様性検定

H_0: 確率ベクトル $X \in \mathbb{S}^{m-1}$は $\mathbb{S}^{m-1}$ 上で一様に分布する

- Rayleigh検定, Random Projection test [Cuesta-Albertos,et al. (2009)]
- $\Rightarrow$ 新たな検定方法の開発