随机矩阵理论的研究（一）：元素独立随机方阵

朱梦

初稿于 2025-06-12，修改于 2025-06-14

1. 奇异值与奇异向量

（一）奇异值分解（SVD）。对于任意矩阵 $\boldsymbol{A} \in \mathbb{C}^{m\times n}$，可分解为：

$$\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^* \tag{1}$$

其中，$\boldsymbol{U} \in \mathbb{C}^{m\times m}$ 表示左奇异向量组成的正交矩阵，对应 $\boldsymbol{A}\boldsymbol{A}^*$ 的特征向量，$\boldsymbol{\Sigma} \in \mathbb{R}^{m\times n}$ 表示对角矩阵，对角线元素为奇异值 $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r$（$r$ 为矩阵的秩），$\boldsymbol{V} \in \mathbb{C}^{n\times n}$ 表示右奇异向量组成的正交矩阵，对应 $\boldsymbol{A}^*\boldsymbol{A}$ 的特征向量。

（二）奇异值的性质。(1）非负性：$\sigma_i \geq 0$。(2）降序排列：$\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r$。

（三）奇异向量的性质。(1）正交归一性：$\boldsymbol{U}$ 的列向量两两正交且模长为 1，$\boldsymbol{V}$ 的列向量两两正交且模长为 1。(2）方向性：左奇异向量指向数据的主成分方向，右奇异向量指向原始特征的主方向。

（四）几何解释。SVD 将矩阵变换分解为旋转（$\boldsymbol{V}^{\mathrm{T}}$）、缩放（$\boldsymbol{\Sigma}$）和再旋转（$\boldsymbol{U}$）三个步骤。奇异值越大，对应方向的“拉伸”效果越显著。

2. 特征值与特征向量

（一）特征值与特征向量的定义。对于方阵 $\boldsymbol{A} \in \mathbb{C}^{n\times n}$，若存在标量 $\lambda \in \mathbb{C}$ 和非零向量 $\boldsymbol{x} \in \boldsymbol{C}^n$，使得：

$$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \tag{2}$$

则称：λ 为 $\boldsymbol{A}$ 的特征值，$\boldsymbol{x}$ 为 $\boldsymbol{A}$ 的对应于 λ 的特征向量。

（二）特征值的代数性质。(1）特征值之和等于矩阵的迹（对角线元素之和）。特征值之积等于矩阵的行列式。(2）不同特征值对应的特征向量线性无关。重根特征值可能对应线性相关特征向量。(3）实对称矩阵的特征值为实数，且不同特征值对应的特征向量正交。

（三）特征值的几何行为。当 λ 为实数，矩阵 $\boldsymbol{A}$ 在特征向量方向上的变换为纯拉伸或压缩，无旋转成分。当 λ 为纯虚数，矩阵 $\boldsymbol{A}$ 在特征向量方向上的变换为纯旋转，无拉伸或压缩。当 λ 为实数，矩阵 $\boldsymbol{A}$ 在特征向量方向上的变换结合了拉伸/压缩与旋转。

3. 谱范数与谱半径

谱范数为矩阵的一种诱导范数，定义为矩阵的最大奇异值：

$$||\boldsymbol{A}||_2 = \sigma_{\max}(\boldsymbol{A}) = \sqrt{\lambda_{\max}(\boldsymbol{A}^*\boldsymbol{A})} \tag{3}$$

其中，$\sigma_{\max}(\boldsymbol{A})$ 表示矩阵 $\boldsymbol{A}$ 的最大奇异值，$\sqrt{\lambda_{\max}(\boldsymbol{A}^*\boldsymbol{A})}$ 表示矩阵 $\boldsymbol{A}^*\boldsymbol{A}$ 的最大特征值。对于线性变换 $\boldsymbol{A}\boldsymbol{x}$，谱范数 $||\boldsymbol{A}||_2$ 表示向量 $\boldsymbol{x}$ 在变换后可能被拉伸的最大倍数，即 $||\boldsymbol{A}\boldsymbol{x}||_2 \leq ||\boldsymbol{A}||_2||\boldsymbol{x}||_2$。

谱半径是矩阵所有特征值的模的最大值：

$$\rho(\boldsymbol{A}) = \max_i |\lambda_i| \tag{4}$$

其中，λ_i 表示是矩阵 $\boldsymbol{A}$ 的特征值。

对于任意复数矩阵 $\boldsymbol{A}$，都有：

$$\rho(\boldsymbol{A}) \leq ||\boldsymbol{A}||_2 \tag{5}$$

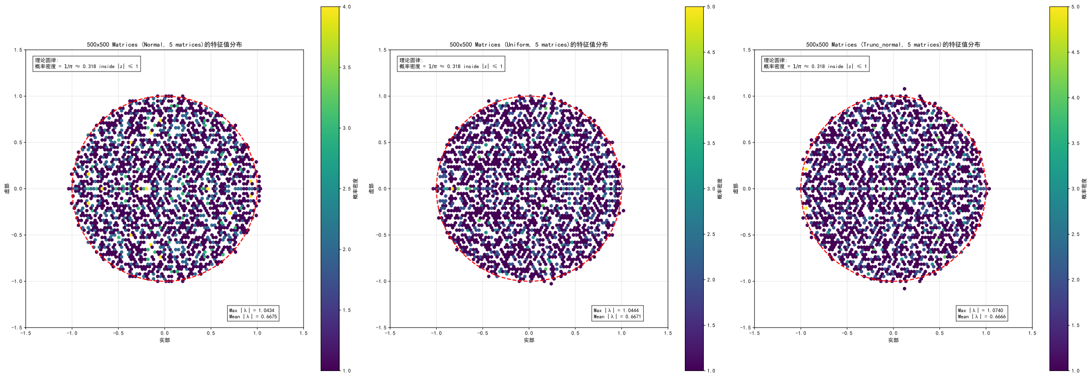
当且仅当矩阵 $\boldsymbol{A}$ 为正规矩阵（即 $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^*$）时，谱半径等于谱范数：

4. 独立元素随机方阵

对于一个随机方阵 $\boldsymbol{A} \in \mathbb{R}^{n\times n}$，其元素独立重复采样于均值为 0，方差为 $1/n$ 的随机分布（随机分布可以是正太分布、均匀分布或截断正太分布）。当 n 很大时，其特征值分布服从圆律（Circular Law），即当 n 很大时，特征值在复平面上均匀分布在以原点为中心、半径为 1 的圆盘内：

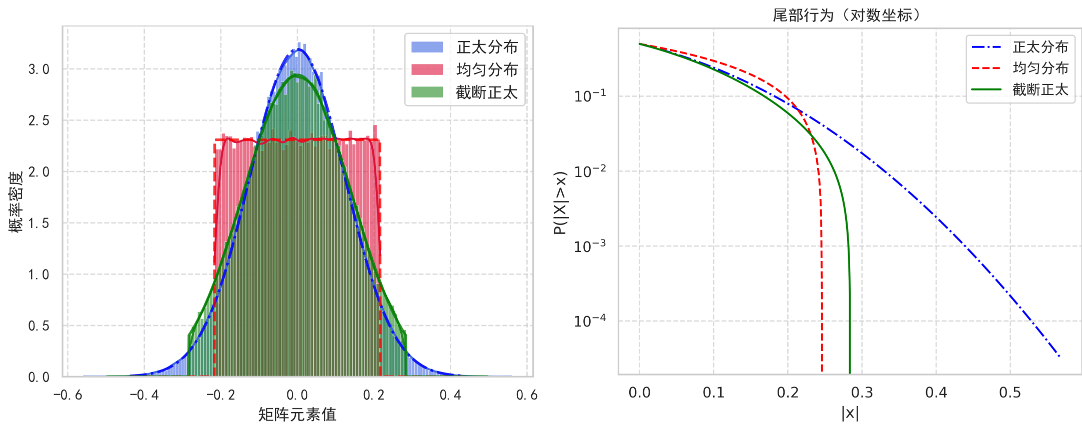
$$|\lambda_i| \leq 1, \lambda_i \in \mathbb{C} \tag{6}$$

$$f_{\text{pdf}}(\lambda_i) = \begin{cases} \frac{1}{\pi}, & \text{如果}|\lambda_i| \leq 1, \\ 0, & \text{其它} \end{cases} \tag{7}$$



(a) 正太分布　　(b) 均匀分布　　(c) 截断正太

图 1　从不同随机分布采样所得到方阵的特征值分布



(a) 直方图　　(b) 尾部行为

图 2　从三类随机分布采样所得到方阵的元素分析及其比较

为什么高维下会有随机向量“近似正交归一化”的直觉？不妨设 $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$

的分量都是从均值为 0，方差为 $1/n$ 的随机分布中独立重复采样出来的，那么有：

$$
\begin{aligned}
<\boldsymbol{u},\boldsymbol{v}> &= \left(n\times\frac{1}{n}\right)\sum_{i=1}^{d}u_iv_i \\
&\approx n\times\mathbb{E}_{u_i\sim f_{\text{PDF}}(x),v_i\sim f_{\text{PDF}}(x)}[u_iv_i] \\
&= n\times\mathbb{E}_{u_i\sim f_{\text{PDF}}(x)}[u_i]\times\mathbb{E}_{v_i\sim f_{\text{PDF}}(x)}[v_i] \\
&= 0
\end{aligned}
\tag{8}
$$

$$
\begin{aligned}
<\boldsymbol{u},\boldsymbol{u}> &= \left(n\times\frac{1}{n}\right)\sum_{i=1}^{n}u_i^2 \\
&\approx n\times\mathbb{E}_{u_i\sim f_{\text{PDF}}(x)}[u_i^2] \\
&= n\times(\mu^2+\sigma^2) \\
&= 1
\end{aligned}
\tag{9}
$$

因此，对任意固定向量 $\boldsymbol{x}$，有：$||\boldsymbol{A}\boldsymbol{x}||_2\approx||\boldsymbol{x}||_2$（等距近似），但这不改变方阵 $\boldsymbol{A}$ 的特征值在复平面内部的分布！

表 1　从三类随机分布采样所得到方阵的比较

随机分布类型	特征值分布	与正交矩阵的近似程度
正太分布	精确圆律（解析可证）	近似最低（元素无界）
均匀分布	渐近圆律（普适性原理）	近似中等（元素有界）
截断正太	渐近圆律（普适性原理）	近似最高（元素有界且接近正太）

```
import numpy as np
from scipy.stats import norm, uniform, truncnorm

# 设置参数
n = 500  # 矩阵维度
var = 1 / n
sigma = (1 / n) ** 0.5
np.random.seed(42)
x = np.random.normal(1, size=n)
I = np.eye(n)


def print_stats(name, data):
    print(f"\n{name}分布统计:")
    print(f"均值: {np.mean(data):.6f} (理论: 0.0)")
```

```
    print(f"方差: {np.var(data):.6f} (目标: {var:.6f})")

    '''数字验证Theta是否近似正交矩阵
    '''
    M = np.einsum("ik, kj->ij", data.T, data)
    err = np.sqrt(np.mean(np.square(M - I)))
    print(f"与正交矩阵的近似程度（越小越好）: {err:.6f} ")

    '''数字验证等距近似
    '''
    z = np.einsum("ij,j->i", data, x)
    x_norm = np.sqrt(np.mean(np.square(x)))
    z_norm = np.sqrt(np.mean(np.square(z)))
    err = abs(z_norm - x_norm)
    print(f"等距近似的程度（越小越好）: {err:.6f} ")


normal_matrix = norm.rvs(loc=0, scale=sigma, size=(n, n))

a = (3 / n) ** 0.5
uniform_matrix = uniform.rvs(loc=-a, scale=2 * a, size=(n, n))

t = 2  # 截断阈值（标准差倍数）
larger_sigma = sigma * 1.1368472343385565
a, b = -t * larger_sigma, t * larger_sigma  # 截断边界
trunc_normal_matrix = truncnorm.rvs((a - 0) / larger_sigma, (b - 0) / larger_sigma,
    loc=0, scale=larger_sigma,
size=(n, n))


print_stats("正态", normal_matrix)
print_stats("均匀", uniform_matrix)
print_stats("截断正态", trunc_normal_matrix)
```

输出结果为:

```
正态分布统计:
均值: 0.000007 (理论: 0.0)
方差: 0.002001 (目标: 0.002000)
与正交矩阵的近似程度（越小越好）: 0.044737
等距近似的程度（越小越好）: 0.034251

均匀分布统计:
均值: 0.000042 (理论: 0.0)
方差: 0.001995 (目标: 0.002000)
与正交矩阵的近似程度（越小越好）: 0.044605
等距近似的程度（越小越好）: 0.008731

截断正态分布统计:
```

```
均值: 0.000042 (理论: 0.0)
方差: 0.002000 (目标: 0.002000)
与正交矩阵的近似程度(越小越好): 0.044687
等距近似的程度(越小越好): 0.029288
```