Random Matrix Theory problems (Work in progress).

Solutions by Gustavo Fuentes

Contents

1 Deterministic Matrices

Problem 1.1. Instability of eigenvalues of non-symmetric matrices

Consider the $N \times N$ square band diagonal matrix $\mathbf{M}_0$ defined by $[\mathbf{M}_0]_{ij} = 2\delta_{i,j-1}$:

$$\mathbf{M}_0 = \begin{bmatrix} 0 & 2 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 2 \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

(a) Show that $\mathbf{M}_0^N = 0$ and so all the eigenvalues of $\mathbf{M}_0$ must be zero. Use a numerical eigenvalue solver for non-symmetric matrices and confirm numerically that this is the case for $N = 100$.

(b) If $\mathbf{O}$ is an orthogonal matrix ($\mathbf{O}\mathbf{O}^\intercal = \mathbf{1}$), $\mathbf{O}\mathbf{M}_0\mathbf{O}^\intercal$ has the same eigenvalues as $\mathbf{M}_0$. Following Exercise 2, generate a random orthogonal matrix $\mathbf{O}$. Numerically find the eigenvalues of $\mathbf{O}\mathbf{M}_0\mathbf{O}^\intercal$. Do you get the same answer as in (a)?

(c) Consider $\mathbf{M}_1$ whose elements are all equal to those of $\mathbf{M}_0$ except for one element in the lower left corner $[\mathbf{M}_1]_{N,1} = (1/2)^{N-1}$. Show that $\mathbf{M}_1 = \mathbf{1}$; more precisely, show that the characteristic polynomial of $\mathbf{M}_1$ is given by $\det(\mathbf{M}_1 - \lambda\mathbf{1}) = \lambda^N - 1$, therefore $\mathbf{M}_1$ has N distinct eigenvalues.

(d) For N greater than about 60, $\mathbf{O}\mathbf{M}_0\mathbf{O}^\intercal$ and $\mathbf{O}\mathbf{M}_1\mathbf{O}^\intercal$ are indistinguishable to machine precision. Compare numerically the eigenvalues of these two rotated matrices.

Solution 1.1. (a) The elements of $[\mathbf{M}_0^N]_{ij}$ for $1 \le i, j \le N$ are given by the following product:

$$[\mathbf{M}_0^N]_{ij} = [\mathbf{M}_0\mathbf{M}_0^N]_{ij} = \sum_k [\mathbf{M}_0]_{ik}[\mathbf{M}_0^{N-1}]_{kj} =$$
$$= \sum_{k_1}\sum_{k_2}[\mathbf{M}_0]_{ik_1}[\mathbf{M}_0]_{k_1k_2}[\mathbf{M}_0^{N-1}]_{k_2j}$$

repeating the product N times, then

$$= \sum_{k_1,k_2,\dots,k_N}^{N} [\mathbf{M}_0]_{ik_1}[\mathbf{M}_0]_{k_1k_2}\cdots[\mathbf{M}_0]_{k_{N-1}k_N}[\mathbf{M}_0]_{k_Nj} =$$
$$= \sum_{k_1=2}^{N}\sum_{k_2=k_1+1}^{N}\cdots\sum_{k_N=k_1+N-1}^{N} 2^N\delta_{i,k_1-1}\delta_{k_1-1,k_2-2}\cdots\delta_{k_{N-1}-N,k_N-N}\delta_{k_N-N+1,j-N}$$

However, in order to obtain a non-zero value $j - N - 1 = i \Rightarrow j = i + N$ and this is out of bounds due to $1 \le j \le N$.

$\therefore [\mathbf{M}_0^N]_{ij} = 0, \forall i, j.$

The eigenvalues are given by:

$$\det(\mathbf{M}_0 - \lambda\mathbf{1}) = 0$$
$$\det\left(\begin{bmatrix} -\lambda & 2 & 0 & \dots & 0 \\ 0 & -\lambda & 2 & \dots & 0 \\ & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 2 \\ 0 & 0 & 0 & \dots & -\lambda \end{bmatrix}\right) = 0$$

This is a triangular matrix, then its determinant equals the product of the diagonal entries

$$\Rightarrow \prod_{k=1}^{N} (-1)^N \lambda_k = 0$$

then all its eigenvalues are $\lambda = 0$.

```
import numpy as np
M = 2*np.eye(100, k=1)
uM, _ = np.linalg.eig(M)
plt.hist(uM, bins='auto',ec = 'Black', color = 'slateblue', density = True)
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')
plt.xlim(-4,4)
plt.show()
```

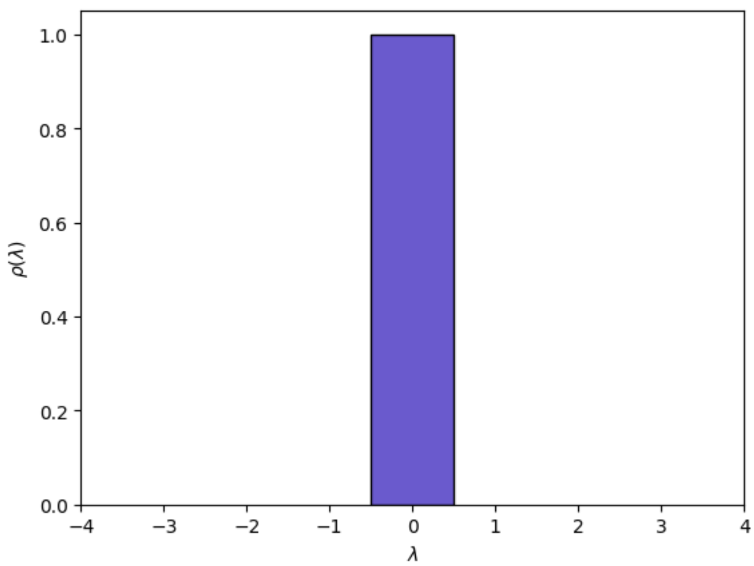

Figure 1: Delta distribution.

(b) The eigenvalues are not the same.

```
O = rmt(N = 100).orthogonal()
M0new = O@M@O.T
unew, _ = np.linalg.eig(M0new)
plt.hist(np.real(unew), bins=30,ec = 'Black', color = 'slateblue', density
                                          = True)
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')
plt.title('$OM_{0}O^{T}$ eigenvalues')
plt.show()
```

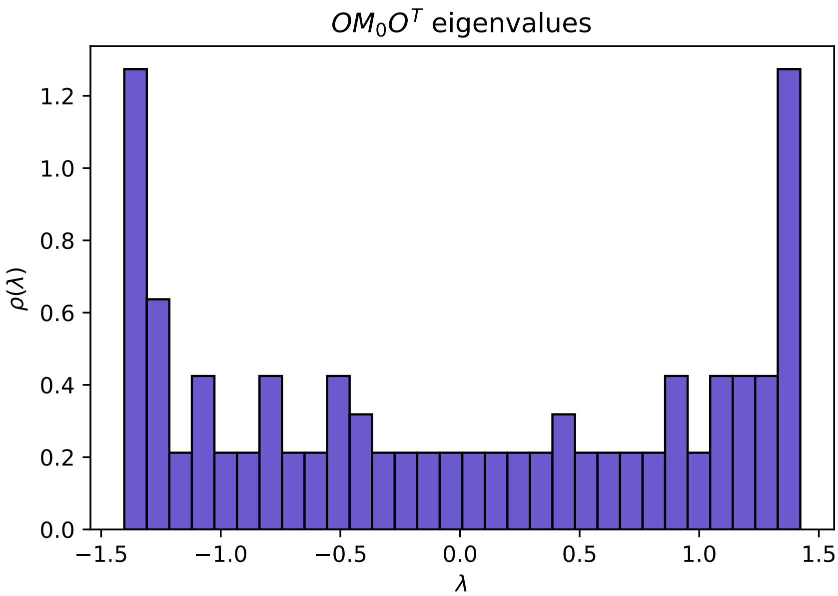

Figure 2: Eigenvalues of the product $\mathbf{OM_oO}^T$

(c) The non-zero elements of $\mathbf{M}_1$ are given by $[\mathbf{M}_1]_{i,i+1} = 2$, $[\mathbf{M}_1]_{N,1} = (1/2)^{N-1}$. The non-zero elements of $[\mathbf{M}_1^N]_{ij}$ are those for $i = j$ and that involves, other wise we have the same case as a) $[\mathbf{M}_1]_{N,1} = (1/2)^{N-1}$, this is:

$$[\mathbf{M}_1^N]_{i,i} = [\mathbf{M}_1]_{i,i+1}[\mathbf{M}_1]_{i+1,i+2}\cdots[\mathbf{M}_1]_{N-1,N}[\mathbf{M}_1]_{N,1}[\mathbf{M}_1]_{1,2}[\mathbf{M}_1]_{2,3}\cdots[\mathbf{M}_1]_{i-1,i} =$$
$$= 2\cdot 2\cdot 2\cdots\frac{1}{2^{N-1}}\cdot 2\cdots 2 = 1$$
$$\therefore \mathbf{M}_1^N = \mathbf{1}$$

The eigenvalues are given by:

$$\det(\mathbf{M}_1 - \lambda\mathbf{1}) = 0$$
$$\det\left(\begin{bmatrix} -\lambda & 2 & 0 & ... & 0 \\ 0 & -\lambda & 2 & \dots & 0 \\ & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & ... & 0 \\ 0 & 0 & 0 & ... & 2 \\ \frac{1}{2^{N-1}} & 0 & 0 & ... & -\lambda \end{bmatrix}\right) = 0$$
$$= -\lambda(-\lambda)^{N-1} - (-1)^N \frac{1}{2^{N-2}} 2^{N-1} = \lambda^N - 1 = 0$$
$$\therefore \lambda_k = \cos\left(\frac{2\pi k}{N}\right)$$
$$\lambda_k = \exp\left(\frac{2\pi i k}{N}\right)$$

(d) For $N = 61$

```
def M1(N):
    M = 2*np.eye(N, k = 1)
    M[N-1,0] = 2**(-1*N+1)
    return M
M0 = 2*np.eye(61, k = 1)
O0 = rmt(N = 61).orthogonal()
M0new = O0@M0@O0.T
u0new, _ = np.linalg.eig(M0new)

M11 = M1(N = 61)
O1 = rmt(N = 61).orthogonal()
M1new = O1@M11@O1.T
u1new, _ = np.linalg.eig(M1new)

plt.figure(figsize = (8,8))
plt.subplot(2,1,1)
plt.hist(np.real(u0new), bins=30,ec = 'Black', color = 'slateblue',
                                            density = True
     ,label = '$OM_{0}O^{T}$')
plt.legend(loc = 'upper center')
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')

plt.subplot(2,1,2)
plt.hist(np.real(u1new), bins=30,ec = 'Black', color = 'red', density =
                                            True
     ,label = '$OM_{1}O^{T}$')
plt.legend(loc = 'upper center')

plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')

plt.show()
```

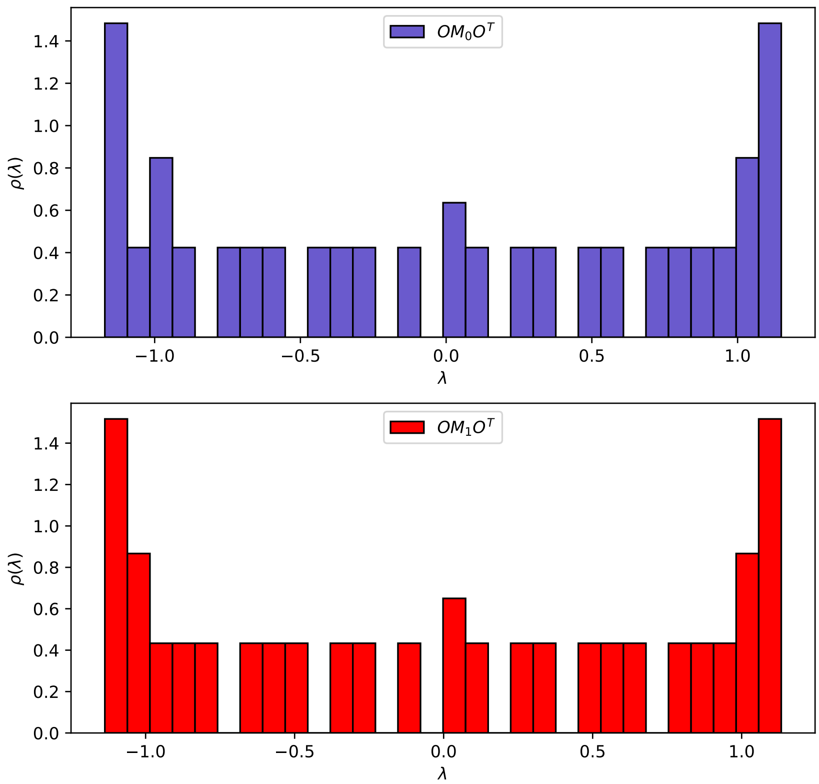

Figure 3: Eigenvalues comparision.

Problem 1.2. Gershgorin and Perron-Frobenius.

Show that the upper bound in

$$\min_j \sum_j \mathbf{A}_{ij} \leq \lambda_{\max} \leq \max_j \sum_j \mathbf{A}_{ij} \tag{1}$$

is a simple consequence of the Gershgorin theorem.

Solution 1.2. By the Gershgorin theorem $\exists \mathcal{D}_i : \lambda_{max} \in \mathcal{D}_i$
Then by the Perron-Frobenius theorem

$$\lambda_{\max} \leq \max_i \sum_j \mathbf{A}_{ij} \leq \sum_{i \neq j} \mathbf{A}_{ij} = R_i$$

due to the elements $\mathbf{A}_{ij} > 0$, then the left side is consequence of the Gershgorin theorem.

Problem 1.3. Sherman-Morrison.

Show that

$$\left(\mathbf{A} + \mathbf{u}\mathbf{v}^T\right)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} \tag{2}$$

is correct by multiplying both sides by $\left(\mathbf{A} + \mathbf{u}\mathbf{v}^T\right)$.

Solution 1.3. Left side:

$$\left(\mathbf{A} + \mathbf{u}\mathbf{v}^T\right)\left(\mathbf{A} - \mathbf{u}\mathbf{v}^T\right)^{-1} = \mathbf{1} \tag{3}$$

Right side:

$$\begin{aligned}
&\left(\mathbf{A} + \mathbf{u}\mathbf{v}^T\right)\left(\mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}\right) = \\
&= \left(\mathbf{A} + \mathbf{u}\mathbf{v}^T\right)\mathbf{A}^{-1} - \left(\mathbf{A} + \mathbf{u}\mathbf{v}^T\right)\frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} = \\
&= \mathbf{A}\mathbf{A}^{-1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} - \frac{\mathbf{A}\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} = \\
&= \mathbf{1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} - \frac{\mathbf{1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}}
\end{aligned}$$

Notice that $1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u} \in \mathbb{R}$, then:

$$\begin{aligned}
&= \mathbf{1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} - \frac{\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} + \left(\mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\right)\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} = \\
&= \mathbf{1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} - \frac{\left(1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}\right)\mathbf{u}\mathbf{v}^T\mathbf{A}^{-1}}{1 + \mathbf{v}^T\mathbf{A}^{-1}\mathbf{u}} = \\
&= \mathbf{1} + \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} - \mathbf{u}\mathbf{v}^T\mathbf{A}^{-1} = \mathbf{1}
\end{aligned} \tag{4}$$

Then left side (3) and right side (4) are equal.

Problem 1.4. Combining Schur and Sherman-Morrison.

The Schur complement, also called inversion by partitioning, relates the blocks of the inverse of a matrix to the inverse of blocks of the original matrix. Let $\mathbf{M}$ be an invertible matrix which we divide in four blocks as

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22} \end{bmatrix} \text{ and } \mathbf{M}^{-1} = \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \tag{5}$$

Where $[\mathbf{M}_{11}] = n \times n$, $[\mathbf{M}_{11}] = n \times n$, $[\mathbf{M}_{12}] = n \times (N-1)$,$[\mathbf{M}_{22}] = (N-1) \times (N-1)$ and $[\mathbf{M}_{22}]$ is invertible. The integer n can take any values from 1 to $N-1$.

For $n = 1$ and any $N > 1$, combine the Shur complement (5) and the Sherman-Morrison to show that:

$$\mathbf{Q}_{22} = (\mathbf{M}_{22})^{-1} + \frac{(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}\mathbf{M}_{12}(\mathbf{M}_{22})^{-1}}{\mathbf{M}_{11} - \mathbf{M}_{12}(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}}$$

Solution 1.4. Using the Sherman-Morrison formula (2) for $\mathbf{M}_{22} - \mathbf{M}_{21}\mathbf{M}_{11}\mathbf{M}_{12}$:

$$(\mathbf{M}_{22} - \mathbf{M}_{21}\mathbf{M}_{11}\mathbf{M}_{12})^{-1} = (\mathbf{M}_{22})^{-1} + (\mathbf{M}_{22})^{-1}\mathbf{M}_{21}\left(\mathbf{M}_{11} - \mathbf{M}_{12}(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}\right)^{-1}\mathbf{M}_{12}(\mathbf{M}_{22})^{-1}$$

due to $\mathbf{M}_{11} = (\mathbf{M}_{11})^{-1}$ is a number.

$$\therefore \mathbf{Q}_{22} = (\mathbf{M}_{22})^{-1} + (\mathbf{M}_{22})^{-1}\mathbf{M}_{21}\left(\mathbf{M}_{11} - \mathbf{M}_{12}(\mathbf{M}_{22})^{-1}\mathbf{M}_{21}\right)^{-1}\mathbf{M}_{12}(\mathbf{M}_{22})^{-1}$$

2 Random Matrices.

Problem 2.1.

(i) Let $\mathbf{M}$ be a random real symmetric orthogonal matrix, that is an $N \times N$ matrix satisfying $\mathbf{M} = \mathbf{M}^\intercal = \mathbf{M}^{-1}$. Show that all the eigenvalues of $\mathbf{M}$ are ± 1.

(ii) Let $\mathbf{X}$ be a Wigner matrix, i.e. an $N \times N$ real symmetric matrix whose diagonal and upper triangular entries are iid Gaussian random numbers with zero mean and variance σ^2/N. You can use $\mathbf{X} = \sigma(\mathbf{H} + \mathbf{H}^\intercal)/\sqrt{2\sigma}$ where $\mathbf{H}$ is a non-symmetric $N \times N$ matrix with iid standard Gaussians.

(iii) The matrix $\mathbf{P}_+$ is defined as $\mathbf{P}_+ = \frac{1}{2}(\mathbf{M} + \mathbf{1})$. Convince yourself that $\mathbf{P}_+$ is the projector onto the eigenspace of $\mathbf{M}$ with eigenvalue $+1$. Explain the effect of the matrix $\mathbf{P}_+$ on eigenvectors of $\mathbf{M}$.

(iv) An easy way to generate a random matrix $\mathbf{M}$ is to generate a Wigner matrix (independent of $\mathbf{X}$), diagonalize it, replace every eigenvalue by its sign and reconstruct the matrix. The procedure does not depend on the σ used for the Wigner.

(v) We consider a matrix $\mathbf{E}$ of the form $\mathbf{E} = \mathbf{M} + \mathbf{X}$. To wit, $\mathbf{E}$ is a noisy version of $\mathbf{E}$. The goal of the following is to understand numerically how the matrix $\mathbf{E}$ is corrupted by the Wigner noise. Using the computer language of your choice, for a large value of N (as large as possible while keeping computing times below one minute), for three interesting values of σ of your choice, do the following numerical analysis.

(a) Plot a histogram of the eigenvalues of $\mathbf{E}$, for a single sample first, and then for many samples (say 100).

(b) From your numerical analysis, in the large N limit, for what values of σ do you expect a non-zero density of eigenvalues near zero.

(c) For every normalized eigenvector $\mathbf{v}_i$ of $\mathbf{E}$ compute the norm of the vector $\mathbf{P}_+\mathbf{v}_i$. For a single sample, do a scatter plot of $|\mathbf{P}_+\mathbf{v}_i|^2$ vs λ_i (its eigenvalue). Turn your scatter plot into an approximate conditional expectation value (using a histogram) including data from many samples.

(d) Build an estimator $\mathbf{\Xi}(\mathbf{E})$ of $\mathbf{M}$ using only data from $\mathbf{E}$. We want to minimize the error $\mathcal{E} = \frac{1}{N}||(\mathbf{\Xi}(\mathbf{E}) - \mathbf{M})||_F^2$ where $||\mathbf{A}||_F^2 = \mathrm{Tr}(\mathbf{A}\mathbf{A}^\intercal)$. Consider first $\mathbf{\Xi_1}(\mathbf{E}) = \mathbf{E}$ and then $\mathbf{\Xi_1}(\mathbf{E}) = 0$. What is the error $\mathcal{E}$ of these two estimators? Try to build an ad-hoc estimator $\mathbf{\Xi}(\mathbf{E})$ that has a lower error $\mathcal{E}$ than these two.

(e) Show numerically that the eigenvalues of $\mathbf{E}$ are not IID. For each sample $\mathbf{E}$ rank its eigenvalues $\lambda_1 < \ldots \lambda_N$. Consider the eigenvalue spacing $s_k = \lambda_k - \lambda_{k-1}$ for eigenvalues in the bulk ($.2N < k < .3N$) and ($.7N < k < .8N$). Make a histogram of $\{s_k\}$ including data from 100 samples. Make 100 pseudo-iid samples: mix eigenvalues for 100 different samples and randomly choose N from the 100N possibilities, do not choose the same eigenvalue twice for a given pseudo-iid sample. For each pseudo-iid sample, compute s_k in the bulk and make a histogram of the values using data from all 100 pseudo-iid samples. original data (not iid).

Solution 2.1. (i) Let $|\psi_i\rangle$ eigenvector of $\mathbf{M}$ with corresponding eigenvalue λ_i, i.e, $\mathbf{M}|\psi_i\rangle =$

$\lambda_i|\psi_i\rangle$, then

$$
\begin{aligned}
&(\mathbf{M}|\psi_i\rangle)^\intercal = \langle\psi_i|\mathbf{M}^\intercal = \mathbf{M}^\intercal\lambda_i^* \\
&\Rightarrow (\mathbf{M}|\psi_j\rangle)^\intercal (\mathbf{M}|\psi_i\rangle) = \langle\psi_j|\mathbf{M}^\intercal\mathbf{M}|\psi_i\rangle = \\
&= \langle\psi_j|\lambda_j^*\lambda_i|\psi_i\rangle = \lambda_j^*\lambda_i\langle\psi_j|\psi_i\rangle = \lambda_j^*\lambda_i\delta_{ij} = \lambda_i^*\lambda_i = 1 \\
&\Rightarrow |\lambda_i| = 1 \\
&\therefore \lambda_i = \pm 1
\end{aligned}
$$

(ii) A class for the problem

```
class rmt(object):
        def __init__(self,N,T = 1):
            self.N = N
            self.T = T

        def wigner(self):
            H = np.random.normal(0, 1, size = (self.N,self.N) )
            return (H + H.T)/np.sqrt(2*self.N)

        def wishart(self):
            H = np.random.normal(0,1, size = (self.N, self.T) )
            return (1/self.T)*H@H.T

        def orthogonal(self):
            H = rmt(self.N).wigner()
            u, v = np.linalg.eig(H)
            U = np.diag(np.sign(u))
            return v@U@v.T

X = rmt(N = 100).wigner()
uX,_ = np.linalg.eig(X)
plt.hist(uX, bins=30,ec = 'Black', color = 'slateblue', density = True
                                            )
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')
plt.title('Wigner semi circle')
plt.show()
```

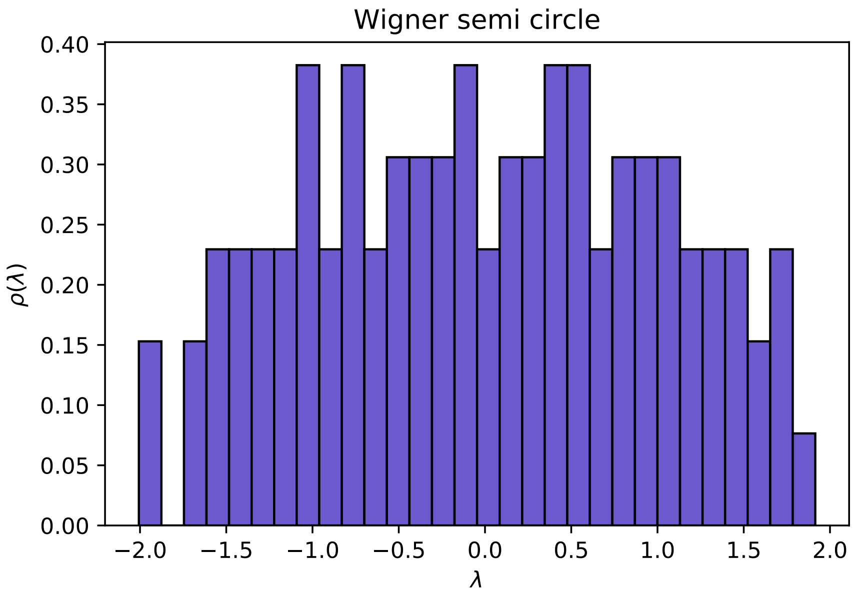

Figure 4: Wigner semi circle law.

(iii) Let $|\psi_i\rangle$ eigenvector of $\mathbf{M}$ with corresponding eigenvalue ± 1, i.e, $\mathbf{M}|\psi_i\rangle = \pm 1|\psi_i\rangle$, then the effect of the $\mathbf{P}_+$ is:

$$\mathbf{P}_+|\psi_i\rangle = \frac{1}{2}(\mathbf{M}+\mathbf{1})|\psi_i\rangle = \frac{1}{2}\left(\mathbf{M}|\psi_i\rangle + \mathbf{1}|\psi_i\rangle\right) = \frac{1}{2}\left(\lambda_i|\psi_i\rangle + |\psi_i\rangle\right)$$
$$\mathbf{P}_+|\psi_i\rangle = \frac{1}{2}(\lambda_i+1)|\psi_i\rangle$$
$$\mathbf{P}_+|\psi_i\rangle = \begin{cases} 1 & \text{if } \lambda_i = 1 \\ 0 & \text{if } \lambda_i = -1 \end{cases}$$

$\mathbf{P}_+$ projects $|\psi_i\rangle$ into itself space if its corresponding eigenvalues is 1 and otherwise project it to zero.

```
def proj(X):
    N = X.shape[0]
    return 0.5*(X + np.identity(N))
P = proj(M)
uP, vP = np.linalg.eig(P)
plt.hist(uP, bins=30,ec = 'Black', color = 'slateblue', density = True
    )
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')
plt.title('$ P$, eigenvalues')
plt.show()
```

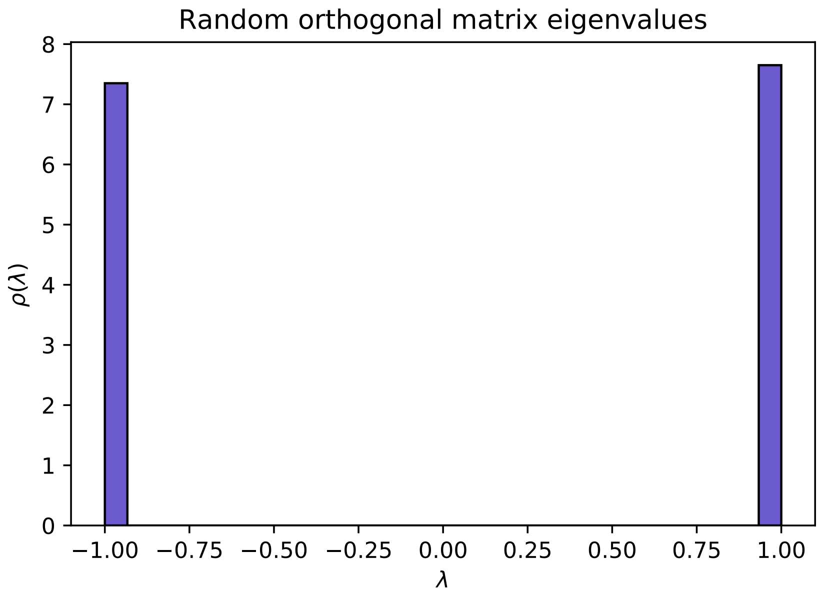

Figure 5: Graphical representation of the problem.

(iv) Orthogonal matrix using the class of ii)

```
M = rmt(N = 100).orthogonal()
uM, _ = np.linalg.eig(M)
plt.hist(uM, bins=30,ec = 'Black', color = 'slateblue', density = True
    )
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')
plt.title('Random orthogonal matrix eigenvalues')
plt.show()
```

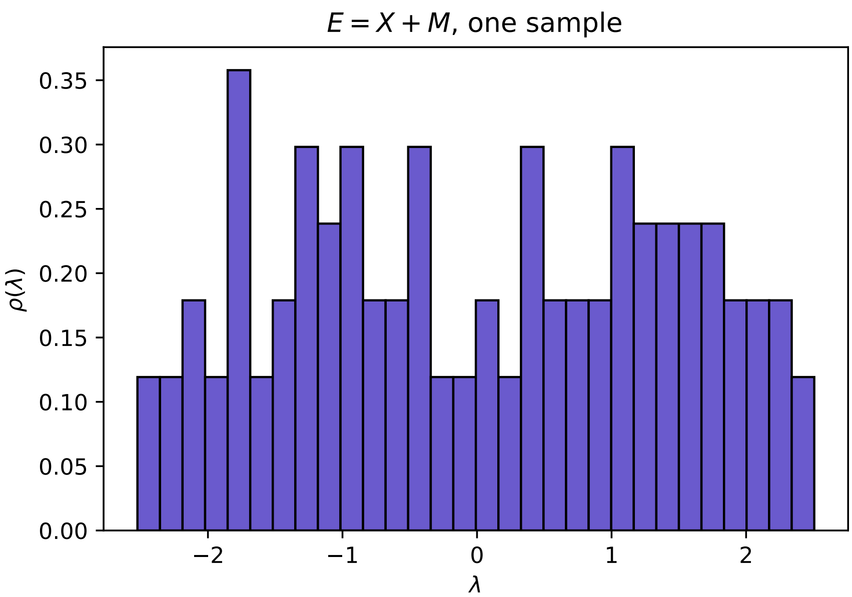

Figure 6: Eigenvalues corresponding to a orthogonal matrix plus a Wigner matrix.

(v) Define $\mathbf{E} = \mathbf{X} + \mathbf{M}$ given the previous indices.

```
E = X + M
```

(a) For one sample

```
uE, vE = np.linalg.eig(E)
plt.hist(uE, bins=30,ec = 'Black', color = 'slateblue', density = True
                                           )
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')
plt.title('$E = X +M $, one sample')
plt.show()
```

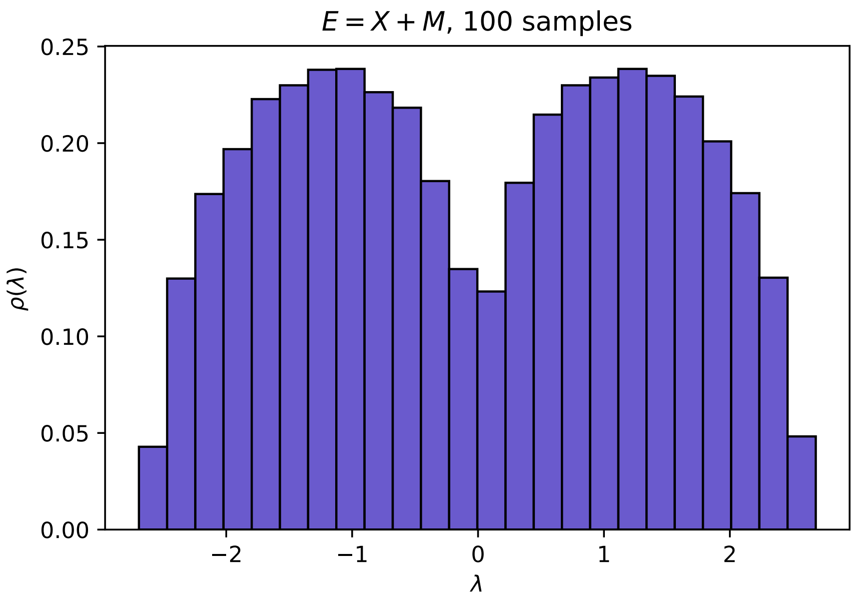

Figure 7: Eigenvalue distribution of a ensemble of 100 members.

For 100 samples

```
uall = []
pall = []
for i in range(100):
    Xi = rmt(N = 100).wigner()
    Mi = rmt(N = 100).orthogonal()
    Ei = Xi + Mi
    Pi = proj(Mi)

    uEi, vEi = np.linalg.eig(Ei)
    pivi = np.abs(np.diag(Pi@vEi))

    uall.append(uEi)
    pall.append(pivi)

uall= np.ravel(uall)

%plt.hist(uall, bins='auto',ec = 'Black', color = 'slateblue', density
                                           = True )
%plt.xlabel('$\lambda$')plt.ylabel('$\\rho(\lambda)$')
```

```
%plt.title('$E = X +M $, 100 samples')
%plt.show()
```

(b) For $\sigma >> 0$

(c) Using pall form the previous index:

```
pall = np.ravel(pall)**2
plt.scatter(pall, uall, color = 'slateblue', alpha = 0.4)
plt.title('$|Pv_{i}|^{2}$ vs $\lambda_{i}$')
plt.ylabel( '$\lambda_{i}$')
plt.xlabel('$|Pv_{i}|^{2}$')
plt.show()
```

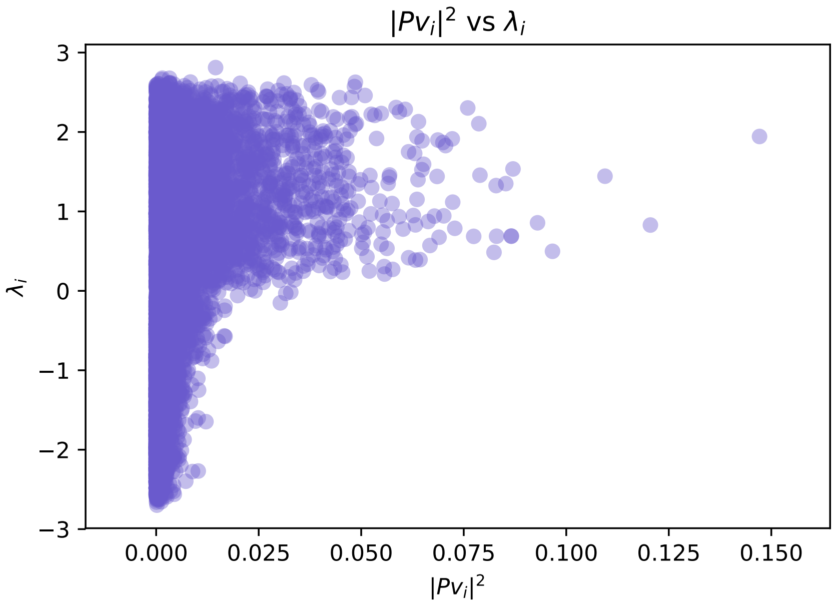

Figure 8: Scatter plot, the negative eigenvalues are projected to zero, while the positives remain the same.

(d) The following function as an estimator.

```
def error( Es, A):
    N = A.shape[0]
    return (1/N)*np.linalg.norm( Es - A, ord = 'fro')**2
Xi1 = error(E,M)
Xi0 = error(np.zeros_like(M), M)
Xieval = error(E/2, M)
```

$Xi1 = 0.98$, $Xi0 = 1$, $Xieval = 0.50$, hence $\Xi(\mathbf{E}) = \mathbf{E}/2$ is the better estimator for this case.

(e) For 100 samples

```
new = np.sort(uall.reshape(100,100),axis = 1)
s1 = new[:,19:29][:,1:]-new[:,19:29][:,:-1]
s2 = new[:,69:79][:,1:]-new[:,69:79][:,:-1]
sk = np.concatenate([s1.reshape(900), s2.reshape(900)], axis = 0)
plt.hist(sk,bins = 'auto', ec = 'black', color = 'slateblue',density =
                                                    True)
plt.xlabel('$\lambda$')
```

```
plt.ylabel('$\\rho(\lambda)$')
plt.show()
```

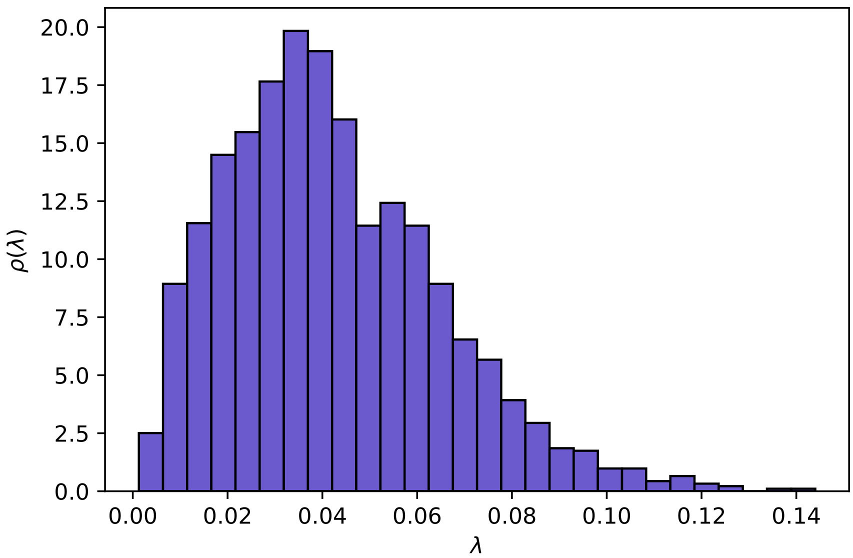

Figure 9: Spacing distribution, similar to the Wigner surmise, the eigenvalues in this case are iid.

```
 np.random.shuffle(uall)
new2 = np.sort(uall.reshape(100,100),axis = 1)
s1 = new2[:,19:29][:,1:]-new2[:,19:29][:,:-1]
s2 = new2[:,69:79][:,1:]-new2[:,69:79][:,:-1]
sk2 = np.concatenate([s1.reshape(900), s2.reshape(900)], axis = 0)
plt.hist(sk2,bins = 'auto', ec = 'black', color = 'slateblue',density =
                                             True)
plt.xlabel('$\lambda$')
plt.ylabel('$\\rho(\lambda)$')
plt.show()
```

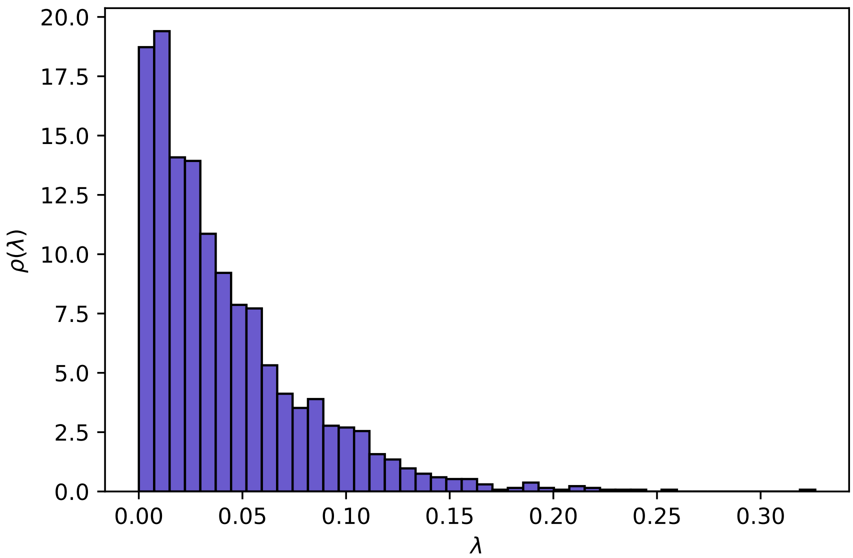

Figure 10: Spacing distribution similar to the Poisson distribution, in this case the eigenvalues are not idd.

3 Wigner ensemble and Semi-circle law.

Problem 3.1. Stieltjes transform for shifted and scaled matrices.

Let $\mathbf{A}$ be a random matrix drawn from a well-behaved ensemble with Stieltjes transform $g(z)$. What are the Stieltjes transforms of the random matrices $\alpha\mathbf{A}$ and $\mathcal{A} + \beta\mathbf{1}$ where α and β are non-zero real numbers and $\mathbf{1}$ the identity matrix?

Solution 3.1. This is equivalent to $g_{\alpha\mathbf{A}+\beta\mathbf{1}}(z)$, by definition:

$$g_{\alpha\mathbf{A}+\beta\mathbf{1}}(z) = \frac{1}{N}\mathrm{Tr}\left(\frac{1}{z\mathbf{1} - \alpha\mathbf{A} - \beta\mathbf{1}}\right) = \frac{1}{N}\mathrm{Tr}\left(\frac{1}{(z-\beta)\mathbf{1} - \alpha\mathbf{A}}\right) =$$

$$= \frac{1}{\alpha N}\mathrm{Tr}\left(\frac{1}{\left(\frac{z-\beta}{\alpha}\right)\mathbf{1} - \mathbf{A}}\right) = \frac{1}{\alpha} g_{\mathbf{A}}\left(\frac{z-\beta}{\alpha}\right)$$

$$\therefore g_{\alpha\mathbf{A}+\beta\mathbf{1}}(z) = \alpha^{-1} g_{\mathbf{A}}\left(\frac{z-\beta}{\alpha}\right)$$

Problem 3.2. Finite N approximation and small imaginary part.

$\mathrm{Im} g_N(x - i\eta)/\pi$ is a good approximation to $\rho(x)$ for a small positive η where $g_N(x)$ is the sample Stieltjes transform ($g_N(z) = (1/N)\sum_k 1/(z - \lambda_k)$). Numerically generate a Wigner matrix of size N and $\sigma^2 = 1$.

(a) For three values of η, $\left\{1/N, 1, 1/\sqrt{N}\right\}$ plot $\mathrm{Im} g_N(x - i\eta)/\pi$ and the theoretical $\rho(x)$ on the same plot for x between -3 and 3.

(b) Compute the error as a function of η where the error is $(\rho(x) - \mathrm{Im} g_N(x - i\eta)/\pi)^2$ summed for all values of x between -3 and 3 spaced by intervals of 0.01. Plot this error for η between $1/N$ and 1. You should see that $1/\sqrt{N}$ is very close to the minimum of this function.

Solution 3.2. On my github.

Problem 3.3. Stieltjes transform.

A large random matrix has moments $\tau(\mathbf{A}^k) = 1/k$

(a) Write the Taylor series of $g(z)$ around infinity

(b) Sum the series to get a simple expression for $g(z)$.

(c) Where are the singularities of $g(z)$ on the real axis?

(d) Redo all the above steps for a matrix whose odd moments are zero and even moments are $\tau(\mathbf{A}^{2k}) = 1$.

Solution 3.3. (a) The Taylor series of $g(z)$ is given by

$$g(z) = \sum_{k=0}^{\infty} \frac{m_k}{z^{k+1}}$$

Where m^k are the moments of the random matrix $\mathbf{A}$, in this case:

$$g(z) = \sum_{k=0}^{\infty} \frac{m_k}{z^{k+1}} = \sum_{k=0}^{\infty} \frac{1}{z^{k+1}}\tau(\mathbf{A}^k) = \sum_{k=0}^{\infty} \frac{1}{z^{k+1}}\frac{1}{k}$$

We have a singularity for $k = 1$ then

$$g(z) = \sum_{k=1}^{\infty} \frac{1}{z^{k+1}} \frac{1}{k} = -\frac{1}{z} \sum_{k=1}^{\infty} -\frac{1}{kz^k}$$

(b) Remark, the Taylor series of $\log(x+1)$ is

$$\log(x+1) = \sum_{k=1}^{\infty} (-1)^{n+1} \frac{x^k}{k}$$

if $x = -\frac{1}{z}$ then

$$\log\left(\left(-\frac{1}{z}\right) + 1\right) = \sum_{k=1}^{\infty} (-1)^{n+1} \frac{(-1)}{kz^k} = \sum_{k=1}^{\infty} (-1)^{k+1+k} \frac{1}{kz^k}$$

$$\therefore \log\left(1 - \frac{1}{z}\right) = \sum_{k=1}^{\infty} -\frac{1}{kz^k}$$

Then the Stieltjes transform is:

$$g(z) = -\frac{1}{z} \log\left(1 - \frac{1}{z}\right)$$

(c) if $z = x - i\eta$ then the singularities on the real axis are $x = 0$ and $x = 1$, also the function is not defined when $x < 1$.

(d) In this case

$$\tau(\mathbf{A}^k) = \begin{cases} 1 & \text{if } k \text{ is even} \\ 0 & \text{if } k \text{ is odd} \end{cases}$$

Then the Stieltjes transform is

$$g(z) = \sum_{k=0}^{\infty} \frac{1}{z^{k+1}} \tau(\mathbf{A}^k) = \sum_{k=0}^{\infty} \frac{1}{z^{2k+1}} \tag{6}$$

but this is the geometric series,

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}$$

for $a = 1$,$r = 1/z$

$$g(z) = \frac{1}{z} \frac{1}{1 - \frac{1}{z^2}}$$

$$\therefore g(z) = \frac{z}{z^2 - 1}$$

with singularities on the real axis when $x = \pm 1$.