Room 517
30 copies
95%
dark
Thanks!

Lecture 2

Before we begin studying the intrinsic properties of random matrices, a few more words on what it means to model a system by RMT. We mean the following: Suppose we are investigating some quantities $\{a_k\}$ in a neighborhood of some point A, say. The a_k's are to be compared with the eigenvalues $\{\lambda_k\}$, in a neighborhood of some point Λ, of a matrix taken from some random matrix ensemble. If the statistics of the a_k's, appropriately centered and scaled,

$$a_k \longmapsto \tilde{a}_k = (a_k - A)\gamma_a$$

some appropriate scaling factor

are described by the statistics of the λ_k's, appropriately centered and scaled

$$\lambda_n \mapsto \tilde{\lambda}_n = (\lambda_n - \Lambda)\,\gamma_\lambda$$

some appropriate scaling factor

then we say that the a_n's are modeled by <u>random matrix theory</u>

<u>Intrinsic Theory</u> : RMT

We will restrict our analysis to ensembles of $N \times N$ Hermitian matrices $\{M = (M_{ij})\}$, $M = M^*$. There are many other ensembles which are of interest, for example ensembles of unitary matrices such as the Circular Unitary Ensemble (CUE) endowed with Haar measure on the Unitary Group. Or COE, the ensemble of (real) orthogonal matrices, also endowed with Haar measure, now on the (real) orthogonal group.

Or sample covariance matrices of the form

$$M = XX^*$$

where X is an $N \times P$ rectangular matrix, whose ($N \times 1$) columns $\vec{x}_1, \ldots, \vec{x}_P$ are identically distributed (Gaussian) samples with mean $\vec{\mu}$ and covariance Σ. Each of these ensembles are useful in various applications (sample covariance matrices are at the heart of what is called Principal Component Analysis, a key tool in statistics and mathematical finance). Although we will not analyze these ensembles (see the various recommended texts & the ref's they contain), the analysis of Hermitian ensembles provides a model and a guide for all random matrix ensembles.

We follow refs (2) and (3) for much of what follows. There are <u>three</u> kinds of Hermitian matrix ensembles which are of interest (Dyson's "threefold way"; see [Mehta]): these consists of

(22.1) $N \times N$ Hermitian matrices $M = (M_{ij}) = M^*$

(22.2) $N \times N$ real Hermitian (ie real symmetric) matrices

$$M = (M_{ij}) = \bar{M} = M^T$$

(22.3) $2N \times 2N$ Hermitian self-dual matrices

$$M = M_{ij} = M^* = J M^T J^+$$

where $J = \mathrm{diag}(\sigma, \sigma, \dots, \sigma)$, $\sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

For all three classes of ensembles, the probability distribution on the matrices is given by

(23.1) $$P_N(M)\,dM = \frac{1}{Z_N} F(M)\,dM$$

where dM is Lebesgue measure on the <u>algebraically independent</u> entries of M, $F(M)$ is a convergence factor to ensure that $P_N(M)\,dM$ is a (finite) probability measure, and

$$Z_N = \int F(M)\,dM$$

is the normalization factor (sometimes called the <u>partition function</u>). (Discrete measures on the matrices, eg $M_{ij} = 1$ or -1 with equal probability are also of interest, but we will not consider them)

Now $N \times N$ Hermitian matrices $M_{jk} = M^R_{jk} + i M^I_{jk}$ $= \overline{M}_{kj}$ depend on $N + 2 \cdot \frac{N(N-1)}{2} = N^2$

and

(24.1) $$dM = \prod_{k=1}^{N} dM_{kk} \prod_{1\le k<j\le N} dM_{kj}^{R} \prod_{1\le k<j\le N} dM_{kj}^{I}$$

For (real) symmetric $N \times N$ matrices $M_{kj} = M_{jk}$, the matrices depend on $N + \frac{N(N-1)}{2} = \frac{N(N+1)}{2}$ real variables and

(24.2) $$dM = \prod_{1\le k\le j\le N} dM_{kj}$$

For $2N \times 2N$ Hermitian self dual matrices M, write M in the form of 2×2 blocks $M = (m_{kj})$, Then (exercise: see also (31)), the $\boxed{\text{~~self dual~~}}$ condition $M = M^* \boxed{\cancel{= J M^T J^T}}$ implies that

(24.3) $$m_{kj} = m_{jk}^{*} \qquad 1 \le j,k \le N$$

and the condition $M^* = -J M^T J$ implies

(24.4) $$\bar{m}_{jk} = -\sigma\, m_{jk}\, \sigma \qquad 1 \le j,k \le N$$

From (24.4) we learn that for

$k \leq j$

(25.1) $$m_{jk} = \alpha_{jk} I + \beta_{jk} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + \gamma_{jk} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \delta_{jk} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

when α_{jk}, β_{jk}, γ_{jk} and δ_{jk} are real

For $k > j$, $m_{jk} = m_{kj}^{+}$ by (24.3).

And for $k = j$, by (25.1) (24.3), $m_{kk} = \alpha_{kk} I$, α_{kk} real. Relation (25.1) shows that the m_{jk}'s are <u>real quaternions</u>. Thus Hermitian self-dual matrices M have 2×2 block structure $M = (m_{ij})$ where the m_{ij}'s are real quaternions. and hence depend on

$$N + 4 \frac{N(N-1)}{2} = 2N^2 - N$$

real variables.

Hence

$$dM = \prod_{k=1}^{N} d\alpha_{kk} \prod_{N \geq k > j \geq 1} d\alpha_{jk}\, d\beta_{jk}\, d\gamma_{jk}\, d\delta_{jk} \qquad (26.1)$$

Historically two kinds of Hermitian ensembles are of interest: those that are invariant under a natural conjugation of the matrices, and those where the algebraically independent entries of the matrices are also statistically independent. The first class of ensembles are called <u>invariant ensembles</u>: the second class are called <u>Wigner ensembles</u>.

We first consider invariant ensembles.

RMT as a subject goes back to the work of statisticians in the 1920's, but the subject was introduced into physics only in the 1950's by Wigner who was interested in the analysis of scattering resonances in neutron scattering theory (the first bookend mentioned above). Such resonances of course reflect precise properties of the Hamiltonian H which describes the neutron–heavy atom system, but the # of degrees of freedom are so large that one could not hope to solve the system for the resonances even numerically. What was needed was a model for the resonances and, taking into account various experimental observations, Wigner was led to positing random matrices as a model for H. Now, as a model for a physical system, a matrix M, say, has no intrinsic meaning: if one changes the basis, the matrix changes by conjugation $M \to M' = UMU^*$.

For this reason Wigner singled out ensembles that were invariant under conjugation by (appropriate) classes of matrices: These are what we called above the invariant ensembles. On the other hand, Wigner ensembles, where the entries are independent, are appropriate for statistical / data type problems, for example, where the matrices have intrinsic meaning (for more info see [Meh] [@]).

Invariance for the Hermitian ensembles (22.1) (22.2) (22.3) means the following:

For (22.1), $P_N(M)\,dM$ must be invariant under the conjugation $M \to UMU^*$ for all unitary matrices U.

For (22.2), $P_N(M)\,dM$ must be invariant under conjugation $M \to UMU^T$ for all (real) orthogonal matrices U

For (22.3) $P_N(M)\,dM$ must be invariant under the conjugation $M \to UMU^*$ for all unitary and symplectic matrices U, i.e. $UU^* = I$ and $UJU^T = J$. Note that if M is a self-dual Hermitian matrix, then so is $M' = UMU^*$ for any unitary/symplectic matrix U. Indeed, M' is clearly Hermitian, so we only need to show that $M'^* = JM'^TJ^T$

But this is a simple exercise.

We first consider invariance for $N \times N$ Hermitian matrices.

For any Hermitian matrix M let

$$\vec{M} = (M_{11}, \ldots, M_{NN}, M_{12}^R, M_{12}^I, M_{13}^R, M_{13}^I, \ldots, M_{N-1,N}^R, M_{N-1,N}^I) \in \mathbb{R}^{N^2}$$

In order to show that

(30.0) $$dM = dM'$$

where $M' = U M U^*$ we must show that $\left| \det \frac{\partial \tilde{M}'}{\partial \tilde{M}} \right| = 1$. But clearly $\operatorname{tr} M^2 = \operatorname{tr} M'^2$

i.e. $\sum M_{jk} M_{kj} = \sum M'_{jk} M'_{kj}$ i.e.

$$\sum_j M_{jj}^2 + 2\sum_{j<k} |M_{jk}|^2 = \sum_j M'^2_{jj} + 2\sum_{j<k} |M'_{jk}|^2$$

or

(30.1) $$\sum_j M_{jj}^2 + 2\sum_{j<k} (M_{jk}^R)^2 + 2\sum_{j<k} (M_{jk}^I)^2$$

$$= \sum_j M'^2_{jj} + 2\sum_{j<k} (M'^R_{jk})^2 + 2\sum_{j<k} (M'^I_{jk})^2$$

In other words if D is the $N^2 \times N^2$ diagonal matrix $D = \operatorname{diag}(1, \ldots, 1, 2, \ldots, 2)$ (N 1's)

then (30.1) states

(30.2) $$(\tilde{M}, D\tilde{M}) = (\tilde{M}', D\tilde{M}')$$

Thus if we write $\tilde{M}' = T\tilde{M}$ for some $N^2 \times N^2$

matrix T, then (30.2) shows that T is orthogonal wrt the inner product induced by D on $\mathbb{R}^{N^2}$ is.

$$T^T D T = D$$

Thus $(\det T)^2 = \det T^T \det T = 1$ & hence

$\det \left|\frac{\partial \tilde{M}'}{\partial \tilde{M}}\right| = \det T = 1$ as desired.

We conclude that $P_N(M)\,dM$ is invariant if and only if

(31.1)
$$\frac{F(M)}{\int F(M)\,dM} = \frac{F(M')}{\int F(M')\,dM'}$$

for all unitary U, $M' = UMU^*$ and all Hermitian M. Setting $M=0$ in (31.1) we see that in fact

$$F(M) = F(M') = F(UMU^*)$$

We will <u>always</u> assume in these lectures that $F(M)$ is of the form

(32.1) $$F(M) = \exp(-\operatorname{tr} Q(M))$$

where $Q(x)$ is a real valued function on $\mathbb{R}$ which grows sufficiently rapidly as $|x| \to \infty$ so that

$$\int e^{-\operatorname{tr} Q(M)}\, dM < \infty .$$

Here $Q(M)$ is defined by the spectral calculus for M i.e. if $M = U \Lambda U^*$, $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_N)$ is the spectral representation for M, U unitary.

Then $$Q(M) = U\, Q(\Lambda)\, U^* = U \begin{pmatrix} Q(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & Q(\lambda_N) \end{pmatrix} U^*$$

For example, if $Q(x) = x^2$, then

$$F(M) = \exp(-\operatorname{tr} M^2)$$

Clearly $F(M) = F(UMU^*)$

This gives rise to the ~~so-called~~ Gaussian Unitary Ensemble (GUE) mentioned in the first lecture,

(32.2) $$P_N(M)\, dM = \frac{1}{Z_N} e^{-\operatorname{tr} M^2} dM .$$

The ensembles of Hermitian matrices with prob. distribution

(33.1) $$P_N(M)\,dM = \frac{1}{Z_N} e^{-\operatorname{tr} Q(M)}\,dM$$

are called the unitary ensembles (UE's).

Which Q is the "right" Q to choose: The remarkable fact is that, on the right scale (as $N\to\infty$), it does not matter! In other words, on the right scale as $N\to\infty$, one has precisely the same fluctuation statistics for the eigenvalues, independent of the particular choice one makes for Q. This phenomenon is known as universality: proving universality is one of the chief tasks in the intrinsic theory of RMT.

For Q, we may choose, for example,

any function $Q(x)$ such that

(34.1) $$Q(x) \geq c x^2 + d$$

for some constants c, d. For then

$\mathrm{tr}\, Q(M) \geq c\, \mathrm{tr}\, M^2 + Nd$ and by (30.1)

$$e^{-\mathrm{tr}\, Q(M)} dM \leq e^{-Nd} e^{-\mathrm{tr}\, M^2} dM$$

$$= e^{-Nd} e^{-M_{11}^2} dM_{11} \cdots e^{-M_{NN}^2} dM_{NN}$$

$$e^{-2|M_{12}^R|^2} dM_{12}^R \; e^{-2|M_{12}^I|^2} dM_{12}^I$$

$$\cdots e^{-2|M_{N-1,N}^R|^2} dM_{N-1,N}^R \; e^{-2|M_{N-1,N}^I|^2} dM_{N-1,N}^I$$

and the RHS clearly has finite integral. In fact

$e^{-\mathrm{tr}\, Q(M)} dM$ clearly has finite moments

(34.2) $$\int |M_{ij}|^k e^{-\mathrm{tr}\, Q(M)} dM < \infty \quad \text{for all } i, j$$

and all $k \geq 0$.

In particular, we can take any polynomial

(34.3) $$Q(x) = \gamma x^{2q} + \cdots \quad , \gamma > 0.$$

As we will see, any $Q(x)$ st

(35.1) $$\frac{Q(x)}{\log|x|} \to +\infty \quad \text{as} \quad |x| \to \infty$$

gives rise to an ensemble with finite moments. (35)

We leave it as an <u>exercise</u> to show that again invariance under orthog. conjugation, and under unitary / symplectic conjugation, for (22.2) and (22.3) respectively, requires the prob. distribution to have the form

$$P_N(M)\,dM = \frac{1}{Z_N} F(M)\,dM$$

and again we always assume that $F(M)$ is of the form $-\operatorname{tr} Q(M)$ for Q as above; thus

$$P_N(M)\,dM = \frac{e^{-\operatorname{tr} Q(M)}}{Z_N}\,dM$$

respectively

One calls such ensembles <u>Orthogonal Ensembles (OE's)</u> / <u>Symplectic Ensembles (SE's)</u>.

If we choose $Q(x) = x^2$ for (22.2) we obtain the Gaussian Orthogonal Ensemble (GOE) and for (22.3) we obtain the Gaussian Symplectic Ensemble (GSE). Note that the only invariant ensembles that are also Wigner ensembles are GUE, GOE and GSE (why?).

In order to compute key statistics, such as the probability that there are no eigenvalues in a gap (a,b) (the so-called gap probability), or the 2-point correlation function, etc., it is useful to use the spectral theorem for M,

$$M = U \Lambda U^*$$

as a change of variables

$$M \mapsto (\Lambda, U)$$

$$dM \to \left| \frac{\partial(M)}{\partial(\Lambda, U)} \right| d\Lambda \, dU$$

It is instructive to consider first the case where $M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ is a real symmetric matrix

Then

$$M = U \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} U^T$$

where $U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, $0 \le \theta < 2\pi$.

Then

$$a = \lambda_1 \cos^2\theta + \lambda_2 \sin^2\theta$$
$$b = (\lambda_1 - \lambda_2)\cos\theta \sin\theta$$
$$c = \lambda_1 \sin^2\theta + \lambda_2 \cos^2\theta$$

and hence

$$\frac{\partial(a,b,c)}{\partial(\lambda_1,\lambda_2,\theta)} = \begin{pmatrix} \cos^2\theta & \sin^2\theta & -2(\lambda_1-\lambda_2)\cos\theta\sin\theta \\ \cos\theta\sin\theta & -\cos\theta\sin\theta & (\lambda_1-\lambda_2)(\cos^2\theta - \sin^2\theta) \\ \sin^2\theta & \cos^2\theta & \lambda_1 \sin^2\theta + \lambda_2\cos^2\theta \end{pmatrix}$$

Thus

(37.1) $$\left| \det\left(\frac{\partial(a,b,c)}{\partial(\lambda_1,\lambda_2,\theta)} \right) \right| = |\lambda_1 - \lambda_2|\, f_1(\theta)$$

where

$$f_1(\theta) = \left| \det \begin{pmatrix} \cos^2\theta & \sin^2\theta & -2\sin 2\theta \\ \frac{1}{2}\sin 2\theta & -\cos\theta\sin\theta & \cos 2\theta \\ \sin^2\theta & \cos^2\theta & \sin 2\theta \end{pmatrix} \right| = 1 > 0$$

If M is 2×2 Hermitian we find similarly (exercise)

(37.2) $$\left| \det\left(\frac{\partial(M)}{\partial(\Lambda, U)} \right) \right| = |\lambda_1 - \lambda_2|^2 f_2(U), \quad f_2(U) > 0$$

and if M is a 4×4 Hermitian self-dual matrix

(38.1) $$\left|\det \frac{\partial M}{\partial(\Lambda, U)}\right| = (\lambda_1-\lambda_2)^4 f_4(U), \quad f_4(U)>0$$

Thus for invariant ensembles, we see (at least for these low-dimensional examples) the eigenvalues and eigenvectors are <u>statistically independent</u>

(38.2) $$e^{-\operatorname{tr} Q(M)}\, dM = e^{-\sum Q(\lambda_i)} \prod_{i<j} |\lambda_i-\lambda_j|^\beta \, d\lambda \, f_\beta(U)\, dU$$

where

(38.3)
$\beta = 1$ for orthogonal ensembles
$\beta = 2$ for unitary ensembles
$\beta = 4$ for symplectic ensembles.

In the well-known analogy between RMT and statistical mechanics (see [Mehta]) β corresponds to an inverse temperature.

We will prove (38.2) (for general M) for orthogonal ensembles ($\beta=2$) and leave the case $\beta=1$ and $\beta=4$ as <u>exercises</u>

(cf 3).

So suppose $M = \bar{M} = M^T$ is an $N \times N$ symmetric matrix. Then M has a spectral decomposition

$$M = U \Lambda U^T$$

where $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N)$ and U is orthogonal.

However the map

$$M \mapsto (\Lambda, U)$$

is not well-defined for all M. Indeed if $M = I$, then any U will do. However, if the eigenvalues of M are simple, $\lambda_i \neq \lambda_j$ for $i \neq j$, then the colums $u_1, \ldots, u_N$, which are the normalized eigenvectors of U, $M u_j = \lambda_j u_j$, are defined up to multiplication by ± 1.

This means that the map

$$M \mapsto (\Lambda, U)$$

is well-defined as a map from the set of matrices with simple spectrum into

$$\{\lambda_1 < \cdots < \lambda_N\} \times O(N)/\mathcal{H}(N)$$

$\mathcal{H}(N) = \mathrm{diag}(\pm 1, \pm 1, \ldots, \pm 1)$

is well-defined and we can compute its Jacobian. This is the approach followed in (2) (3). Note that a critical element in this approach is to show that the set of symmetric matrices with simple spectrum has full measure i.e. its complement has measure 0. This is necessary to conclude that $M \mapsto (\Lambda, U)$ is a valid change of variables for integration. However we will use a slightly different approach suggested by Olivier Conway, which avoids calculations for the homogeneous space $O(N)/H(N)$.

The key fact is the following:

Z