Random Matrices, Ada Altieri

Wishart ensemble (lectures 9 & 10, Feb. 13th & 15th, 2024)

Contents

This is included in chapters 13 and 14 in the book by Giacomo Livan, Marcel Novaes, Pierpaolo Vivo, Introduction to Random Matrices Theory and Practice, and in chapter 3 in the book by Peter Forrester Log-Gases and Random Matrices. For applications of random matrices to mesoscopic physics, see chapter 6 in the book by Konstantin Efetov Supersymmetry in Disorder and Chaos.

1 Wishart is everywhere

1.1 Statistics

Formally, a Wishart distribution for a positive-definite symmetric matrix valued random variable X is characterized by a number of parameters. First there are the dimensions $N \times N$, second an integer $M \geq N$, and third there is an $N \times N$ positive-definite matrix V. The distribution of X reads

$$\rho[X] = \frac{1}{\Gamma_N(M/2)} 2^{-NM/2} (\det V)^{-M/2} (\det X)^{M-N-1} \mathrm{e}^{-\frac{1}{2}\mathrm{Tr}(V^{-1}X)} \tag{1}$$

where the normalization constant involves the mutlivariate Γ function defined by $\Gamma_N(M/2) = \pi^{N(N-1)/4} \prod_{j=1}^{M} \Gamma\left(\frac{M}{2} - \frac{j-1}{2}\right)$.

This distribution arises quite naturally in several contexts, in spite of its nasty looking expression. It can be view as a generalization of the well-known Γ distribution, for a real positive random variable X with density $p(x) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} \mathrm{e}^{-\beta x}$ $(\alpha, \beta > 0)$. The latter is appropriate to describe the kinetic energy of a gas in thermal equilibrium, starting from the (Gaussian) Maxwell distribution for each components of the velocities of the particles. The Wishart distribution can be viewed as a higher-dimensional generalization of the exponential or Γ distribution.

Beyond the formal construction of probability distributions, think of the following problem. We wish to study N traits of a population comprising $M \geq N$ individuals. For each individual $j = 1, \ldots, M$, let's collect these N traits into a column vector $\vec{y}^{(j)}$ with N coordinates. The data matrix $Y = (\vec{y}^{(1)} | \ldots | \vec{y}^{(M)})$ is an $N \times M$ matrix with entries $y_i^{(j)}$ $(i = 1, \ldots, N$ and $j = 1, \ldots, M)$. The empirical average of trait i is given by

$$\bar{y}_i = \frac{1}{M} \sum_{\ell=1}^{M} y_i^{(\ell)} \tag{2}$$

and an empirical approximation of the covariance matrix W quantifying correlations between two traits i and i' has elements

$$W_{ii'} = \frac{1}{M-1} \sum_{j=1}^{M} ()(y_i^{(j)} - \bar{y}_i)(y_{i'}^{(j)} - \bar{y}_{i'}) \tag{3}$$

and it can of course be read as $W = (Y - \bar{Y})^T (Y - \bar{Y})$. If the y fluctuations are Gaussian then W itself will follow a Wishart distribution. In passing, we note that for a given dataset Y characterized by W, it is customary to perform a principal component analysis: this consists in diagonalizing W. The eigenvector corresponding to the largest eigenvalue corresponds to the direction of trait space where the data are the most scattered. This suggests that it might be of interest to look more carefully into the (possible) universal properties of the spectrum of such matrices.

1.2 Physics

In quantum mechanics, or in optics, one if often led to study problems involving wave transmission and reflection. To characterize the transmitted and reflected waves one usually introduces r and t, which are reflection and transmission coefficients for scalar waves. For vectorial waves (or spinors), r and t are matrices. However, from a physics viewpoint, or interest goes to, say, the transmitted energy or intensity, which leads to define a separate energy-related transmission coefficient T. Not surprisingly if one remembers ones electrodynamics courses, $T \sim |t|^2$, or in matrix form, $T \sim t^\dagger t$. This is interesting because if for some reason the transmission matrix t has random elements, then so T will have, and we are back onto the same problem as before.

As experimentally smaller and smaller systems became accessible (mesoscopic systems at the end of the eighties, and later nanoscopic ones) it appeared that size reduction came with new features. Consider a mesoscopic wire-shaped conductor, and run some electric current throught it. It conductivity (or conductance) depends on the sample under study. Build another mesoscopic wire and you will get another value for the conductance. The Landauer formula tells us that the the conductance is related to the fraction of energy transmitted,

$$G = G_0 T, \ G_0 = \frac{e^2}{\pi \hbar} \tag{4}$$

and because T fluctuates from sample to sample, so does G. It was realized (see Lee and Stone, https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.55.1622)that conductance fluctuations were universal (independent of the details of the material used in the conduction experiment). If this was so, perhaps the full details of the transmission matrix t were not needed. In the smae spirit as Wigner, it was then thought that taking t to be Gaussian distributed could suffice, and sure enough, the outcome of the calculation for T is that, indeed the number obtained by taking t a random matrix nicely matches experimental results. This is an important success of the random matrix theory. And this is enough to boost our motivation to look into the Wishart ensemble!

2 Wishart ensemble

2.1 Construction and Coulomb gas picture

In what follows we will work with Hermitian matrices (the parameter $\beta = 2$) but wherever relevant, for real or quaternionic ensembles, β can be replaced with 1 or 4.

Let H be an $N \times M$ ($M \geq N$) Hermitian matrix whose elements are independent and identically distributed Gaussian entries. Define $W = H^\dagger H$. This is an $N \times N$ Hermitian matrix, and it thus has real eigenvalues. Because W is positive semi-definite, its eigenvalues are nonnegative. We will establish this formula a bit later, but for now, we use the fact that the joint pdf of the entries of W reads, up to some normalization,

$$\rho[W] \sim \mathrm{e}^{-\frac{1}{2}\mathrm{Tr}(W)}(\det W)^{\frac{\beta}{2}(M-N+1)-1} \tag{5}$$

and this is the great novelty of this chapter. Now, the matrix elements of W are obviously correlated, due to the presence of the determinant. Keeping in mind that we still need to justify this form of $\rho[W]$, if we accept it, we immediately get the joint pdf of the eigenvalues $\lambda_1, \ldots, \lambda_N$ of W, which involves the partial Jacobian $\prod_{i<j} |\lambda_i - \lambda_j|^\beta$,

$$\rho(\lambda_1, \ldots, \lambda_N) \sim \mathrm{e}^{-\frac{1}{2}\sum_i \lambda_i} \prod_i \lambda_i^{\alpha\beta/2} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \tag{6}$$

where $\alpha = 1 + M - N - 2/\beta$.

We remember that the two-dimensional Coulomb potential between two charges q and q' a distance r apart is $-\frac{qq'}{2\pi\varepsilon_0}\ln r$ (you have most probably derived this result in the form of the potential energy per unit length of two infinite and parallel linearly charged wires in three-dimensional space). Hence for a set of charges q_i with positions $\mathbf{r}_i$ subjected to an external potential V_{ext}, the corresponding canonical partition function is

$$Z = \int \mathrm{d}^2 r_1 \ldots \mathrm{d}^2 r_N \mathrm{e}^{-\frac{1}{T}\sum_i V_{\text{ext}}(\mathbf{r}_i)} \prod_{i<j} |\mathbf{r}_i - \mathbf{r}_j|^{\frac{q_i q_j}{2T\pi\varepsilon_0}} \tag{7}$$

This shows that for a specific temperature T such that $\beta = \frac{q_i q_j}{2T\pi\varepsilon_0}$, and for a careful choice of the external potential, the eigenvalues of W behave as charged particles in two dimensions. That's a useful physical picture to keep in mind. In our particular case there are some specifics:

- the eigenvalues are confined on the real positive axis
- the external potential is such that $V_{\text{ext}}(x)/T = \frac{x}{2} - \frac{\alpha}{2}\ln x$.

Given the choice of M and N, $\alpha \geq 0$ and the external potential turns out to be a confining potential fighting against the Coulomb repulsion (linear growth at $+\infty$ and $-\ln x$ repulsion at $x \to 0^+$).

It is interesting to note that the name of Laguerre is often attached to the Wishart ensemble. It seems that this is becauseLaguerre polynomials, defined by

$$L_n(x) = \frac{\mathrm{e}^x}{n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^n\mathrm{e}^{-x}),\ L_n^{(\alpha)}(x) = \frac{\mathrm{d}^\alpha L_n(x)}{\mathrm{d}x^\alpha} \tag{8}$$

are orthogonal with respect to exactly the same weight as $\mathrm{e}^{-V_{\mathrm{ext}}(x)/T}$, namely

$$\int_0^{+\infty}\mathrm{d}x L_n^{(\alpha)}(x)L_{n'}^{(\alpha)}(x) = \delta_{nn'}\times\frac{\Gamma(n+\alpha+1)}{n!} \tag{9}$$

These have been seen before in the study of the hydrogen atom. The radial part of the wave function of an excited state behaves as $r^\ell\mathrm{e}^{-r}L_{n+\ell}^{(2\ell+1)}(2r)$ where r is measured in unit of the Bohr radius a_0.

Before we look into the consequences of the particular form of the joint pdf of the eigenvalues, let's take a moment to justify the form of $\rho[W]$ in Eq. (5).

2.2 Chiral variation

No times for this. See Peter Forrester's book.

2.3 Derivation of the ensemble

We start from the all-purpose formula

$$\rho[W] = \langle\delta(W - HH^\dagger)\rangle \tag{10}$$

where the average brackets are over the distribution $\mathrm{e}^{-\frac{1}{2}\mathrm{Tr}(HH^\dagger)}$ of the $N\times M$ complex elements of H:

$$\rho_H[H] = \prod_{i=1}^N\prod_{j=1}^M\frac{1}{2\pi}\mathrm{e}^{-\frac{1}{2}(H_{ij}^R)^2-\frac{1}{2}(H_{ij}^I)^2} = (2\pi)^{-NM}\mathrm{e}^{-\frac{1}{2}\mathrm{Tr}(HH^\dagger)} \tag{11}$$

For $\beta = 2$, which we stick to, and the integration volume is

$$\mathrm{d}H = \prod_{i,j}\mathrm{d}H_{ij}^R\mathrm{d}H_{ij}^I \tag{12}$$

A matrix version of $\delta(k) = \int_{-\infty}^{+\infty}\frac{\mathrm{d}t}{2\pi}\mathrm{e}^{itk}$ is now need. Let K me a Hermitian matrix, then one can check that

$$\delta(K) = \frac{1}{2^N\pi^{N^2}}\int\mathrm{d}T\mathrm{e}^{i\mathrm{Tr}(TK)} \tag{13}$$

where $\mathrm{d}T = \prod_i\mathrm{d}T_i i\prod_{i<j}\mathrm{d}T_{ij}^R\mathrm{d}T_{ij}^I$. Once this is accepted, we can write that

$$\rho[W] = \frac{1}{2^N\pi^{N^2}}(2\pi)^{-NM}\int\mathrm{d}T\mathrm{d}H\mathrm{e}^{-\frac{1}{2}\mathrm{Tr}(HH^\dagger)+i\mathrm{Tr}[T(W-HH^\dagger)]} \tag{14}$$

which, up to a simple change of variables, is the same as

$$\rho[W] = \frac{1}{2^N \pi^{N^2}} (2\pi)^{-NM} 2^{-N^2} \int \mathrm{d}T \mathrm{d}H \mathrm{e}^{-\frac{1}{2}\mathrm{Tr}(HH^\dagger) + \frac{i}{2}\mathrm{Tr}[T(W - HH^\dagger)]} \tag{15}$$

Let s be a complex number and let $\mathrm{d}^2 s = \mathrm{d}s\mathrm{d}s^* = 2\mathrm{d}\Re(s = \mathrm{d}\Im(s)$, then one can show that for $\Im(\mu) > 0$,

$$\int \mathrm{d}^2 s \mathrm{e}^{i\mu|s|^2/2 - iT|s|^2/2} = 4\pi i(\mu - T) \tag{16}$$

on condition that the Fresnel integral $\int \mathrm{d}u \mathrm{e}^{iu^2/2} = \sqrt{2\pi}\mathrm{e}^{i\pi/4}$ is known. We want a mtrix version of this identity, and we thus introduce $\vec{s}$, an N-dimensional vector with complex entries, to establish that

$$\det(\mu \mathbf{1}_{N,N} - T)^{-1} = (4\pi i)^{-N} \int \mathrm{d}^{2N} s \mathrm{e}^{\frac{i\mu}{2}\vec{s}^\dagger \vec{s} - \frac{i}{2}\vec{s}^\dagger T \vec{s}} \tag{17}$$

and the integral is repeated M times with $\vec{s}_1, \ldots, \vec{s}_M$ so as to obtain

$$\det(\mu \mathbf{1}_{N,N} - T)^{-M} = (4\pi i)^{-NM} \int \mathrm{d}^{2NM} s \mathrm{e}^{\frac{i\mu}{2}\sum_{i=1}^{M} \vec{s}_i^\dagger \vec{s}_i - \frac{i}{2}\vec{s}_i^\dagger T \vec{s}_i} \tag{18}$$

We then get back to Eq. (15) in which we set $(\vec{s}_j)_i = H^R_{ij} + iH^I_{ij}$ where $i = 1, \ldots, N$ and $j = 1, \ldots, M$ and use that $\mathrm{d}(\vec{s}_j)_i \mathrm{d}(\vec{s}_j)^*_i = 2\mathrm{d}H^R_{ij}\mathrm{d}H^I_{ij}$ and we arrive at the almost formula

$$\rho[W] = 2^{-N}\pi^{-N^2}(2\pi)^{-NM} 2^{-N^2 - NM} (2\pi i)^{NM} \int \mathrm{d}T \mathrm{e}^{\frac{i}{2}\mathrm{Tr}(TW)} \det(i\mathbf{1}_{N,N} - T)^{-M} \tag{19}$$

The last setp is to use the Ingham-Siegel formula

$$\int \mathrm{d}T \mathrm{e}^{\frac{i}{2}\mathrm{Tr}(TW)} \det(\mu \mathbf{1}_{N,N} - T)^{-M} \propto (\det W)^{M-N} \mathrm{e}^{-\frac{1}{2}\mathrm{Tr}\, W} \tag{20}$$

in which we did not care about the proportionality constant.

3 Eigenvalue distribution

3.1 The Marcenko-Pastur distribution, take 1

Our first goal is to determine the pdf of a single eigenvalue. But we need to think a little before doing so. In the GOE eigenvalues are of order $O(\sqrt{N})$ so that we expect them to be of order $O(N)$ in the Wishart ensemble (basically a Wishart matrix is the square of a standard matrix). This will inspire the scaling form of the pdf of an eigenvalue.

Our starting point is the coulomb gas formulation. The canonical partition function reads

$$Z = \int \mathrm{d}\lambda_1 \ldots \mathrm{d}\lambda_N \mathrm{e}^{-\beta E[\lambda_1, \ldots, \lambda_N]} \tag{21}$$

with

$$E = \frac{1}{2\beta}\sum_i \lambda_i - \frac{\alpha}{2}\ln\lambda_i - \sum_{i<j}\ln|\lambda_i - \lambda_j| \tag{22}$$

where$\alpha = 1+M-N-2/\beta$. We want to work in terms of the rescaled variables $x_i = \lambda_i/(\beta N)$. Then up to an irrelevant prefactor,

$$Z = \int \mathrm{d}x_1 \ldots \mathrm{d}x_N \mathrm{e}^{-\beta E} \tag{23}$$

where

$$E = N\frac{1}{2}\sum_i x_i - \frac{\alpha}{2}\ln x_i - \sum_{i<j}\ln|x_i - x_j| \tag{24}$$

We shall work with large N but we want to keep the ration $N/M = c \leq 1$ fixed. This means that $\alpha \simeq N(1/c-1)$. Next we introduce

$$\hat{\rho}(x) = \frac{1}{N}\sum_i \delta(x - x_i) \tag{25}$$

This is the empirical density of (rescaled) eigenvalues. In the large N limit, $\hat{\rho}$ self averages to the pdf of $x = \lambda/N$, which is exactly what we are after. Note that for $\beta = 2$

$$EN^{-2} = F[\hat{\rho}] = \frac{1}{2}\int \mathrm{d}x(x - (1/c-1)\ln x)\hat{\rho}(x) - \frac{1}{2}\int \mathrm{d}x\mathrm{d}x'\hat{\rho}(x)\hat{\rho}(x')\ln|x - x'| \tag{26}$$

and thus

$$Z = \int \mathrm{d}x_1 \ldots \mathrm{d}x_N \mathrm{e}^{-\beta N^2 F[\hat{\rho}]} \tag{27}$$

This is still not perfect. One more step is needed. We would like to convert the N-fold integral over individual positions into an integral over the realizations of $\hat{\rho}$, because we do not mind losing the information on the identity of the particles. In practice, we need to count how many configurations of the positions achieve a give value of the density profile $\hat{\rho}$. The answer is

$$\int \mathrm{d}x_1 \ldots \mathrm{d}x_N \ldots = \int \mathscr{D}\hat{\rho}\mathrm{e}^{N\int \mathrm{d}x\hat{\rho}\ln\hat{\rho}} \tag{28}$$

While the explicit form is not needed, we expected the scaling in N that appears in that equation, because what we need in the exponential is the entropy of the set of N particles, which is extensive in N. Altogether we thus see that

$$Z_0 = \int \mathscr{D}\hat{\rho}\mathrm{e}^{N\int \mathrm{d}x\hat{\rho}\ln\hat{\rho}}\mathrm{e}^{-\beta N^2 F[\hat{\rho}]} \tag{29}$$

For large N, it is clear that the configuration of $\hat{\rho}$, which we call ρ, that dominates this integral is the one that minimizes F:

$$\left.\frac{\delta F}{\delta\hat{\rho}}\right|_{\hat{\rho}=\rho} = 0 \tag{30}$$

This turns into an integral equation:

$$\frac{1}{2}\left(x-(1/c-1)\ln x\right)=\mathrm{vp}\int \mathrm{d}x'\hat{\rho}(x')\ln|x-x'| \tag{31}$$

which we differentiate with respect to x,

$$\frac{1}{2}\left(1-(1/c-1)/x\right)=\mathrm{vp}\int \mathrm{d}x'\frac{\hat{\rho}(x')}{x-x'} \tag{32}$$

Fortunately, this is a well-known form of equation that appears in the hydrodynamics of the airfoil flow, and this has been extensively studied by Tricomi (Integral Equations). He teaches us that the equation

$$\mathrm{vp}\int_a^b \mathrm{d}x'\frac{f(x')}{x-x'}=g(x) \tag{33}$$

has a solution of the form

$$f(x)=\frac{1}{\pi\sqrt{(x-a)(b-x)}}\left[C-\frac{1}{\pi}\mathrm{vp}\int_a^b \mathrm{d}x' g(x')\frac{\sqrt{(x'-a)(b-x')}}{x-x'}\right] \tag{34}$$

Of course, we still do not know a and b, and we further assume that the support of ρ is a compact interval $[a,b]$, which is motivated by the physical picture of confined particles that prevents the splitting of the fluid into separate clusters.

The solution we find is actually quite simple:

$$\rho(x)=\frac{1}{2\pi x}\sqrt{(x-\zeta_-)(\zeta_+-x)},\quad \zeta_\pm\left(1/\sqrt{c}\pm 1\right)^2 \tag{35}$$

The idea is to leave a and b undetermined, and then to evaluate F at the a and b depenendent expression of ρ, and then to minimise F with respect to a and b. The formula in Eq. (35) is the famous Marcenko-Pastur distribution. It notably differs from the Wigner semi-circle law: it has two soft edges away from $x=0$, and actually there is a steep repulsion close to 0 in case $c\to 1^-$ due to the $1/x$ prefactor. The final result is independent of β as was obvious from the early rescaling.

3.2 The marcenko-Pastur distribution, take 2, using the reolvent

Starting again from the Coulomb gas formulation in Eq. (24) we search for the saddle directly, without discussing why we can omit the integration measure, and we find that

$$\frac{1}{2}+\left(\frac{1}{2}-\frac{1}{2c}\right)\frac{1}{x_i}=\frac{1}{N}\sum_{j\neq i}\frac{1}{x_i-x_j} \tag{36}$$

and we multiply by $\frac{1}{N(z-x_i)}$ and sum ofver i. We now introduce the quantity

$$G_N(z)=\frac{1}{N}\sum_i\frac{1}{z-x_i} \tag{37}$$

known as the resolvent, which can later be used to get $\rho(x)$. For instance

$$\langle f(x_i)\rangle = \frac{1}{2\pi i}\oint \mathrm{d}z G_N(z) f(z) \tag{38}$$

as $N \to +\infty$. But the standard reason for introducing G_N is the following. Consider $\hat{\rho}(x) = \sum_i \delta(x - x_i)$, which we are after in the large N limit (using self averaging), then

$$G_N(z) = \frac{1}{N}\mathrm{Tr}\frac{1}{z\mathbf{1} - W} \tag{39}$$

and given that

$$\lim_{\varepsilon\to 0^\pm} \frac{1}{u \pm i\varepsilon} = \mathrm{vp}\frac{1}{u} \mp i\pi\delta(u) \tag{40}$$

we see that $\rho(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0^+} \Im G_N(x - i\varepsilon)$. This is formula 8.8 in Vivo *et al.*'s book.

We multiply Eq. (36) with $\frac{1}{N}\frac{1}{z-x_i}$ and sum over i:

$$\frac{1}{2}G_N + (1/2 - c^{-1}/2)\frac{1}{z}(K + G_N)) = \frac{1}{2}G_N^2 + O(1/N) \tag{41}$$

where we have used that $\frac{1}{x_i(z-x_i)} = \frac{1}{z}\frac{1}{x_i} + \frac{1}{z}\frac{1}{z-x_i}$ and where $K = \langle 1/x_i\rangle$ is just a number. Setting $\gamma = c^{-1} - 1$ we obtain

$$G_N(z) = \frac{1}{2}\left[\pm\frac{\sqrt{\gamma^2 - 4\gamma K z + z^2 - 2\gamma z}}{z} - \frac{\gamma}{z} + 1\right] \tag{42}$$

and we now write $z = x - i\varepsilon$ and work out the imaginary part as $\varepsilon \to 0^+$ (and x real). We readily find that

$$\rho(x) = \frac{1}{\pi}\lim_{\varepsilon\to 0^+} \Im G_N(x - i\varepsilon) = \frac{1}{2\pi x}\sqrt{(x - \zeta_-)(\zeta_+ - x)} \tag{43}$$

where $\zeta_\pm = \gamma(-2\sqrt{K^2 + K} + 2K + 1)$ are yet undetermined. Finally, suing Mathematica, we see that

$$\int_a^b \mathrm{d}x \frac{\sqrt{(x-a)(b-x)}}{2\pi x} = \frac{1}{4}\left(-2\sqrt{ab} + a + b\right) \tag{44}$$

so that, with $a = \zeta_-$ and $b = \zeta_+$, normalization forces $K = 1/\gamma$. This is consistent with the fact that $K = \int \mathrm{d}x \frac{\hat{\rho}(x)}{x}$. Here we have $a = (1 - 1/\sqrt{c})^2$ and $b = (1 + 1/\sqrt{c})^2$. And the expressions for $\zeta_\pm = (1 \pm 1/\sqrt{c})^2$ are recovered. The plot gives

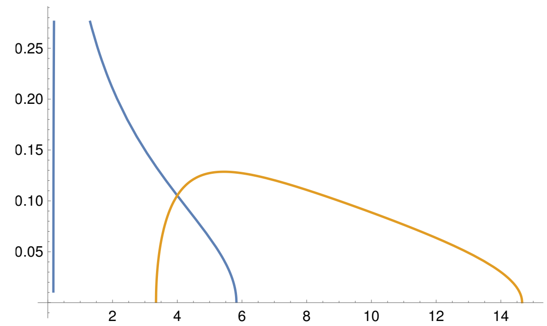

Figure 1: Plot for $c = 1/2$ (blue) and $c = 1/8$ (orange).

3.3 The largest eigenvalue (general strategy)

This is based on P. Vivo, Large deviations of the maximum eigenvalue in Wishart random matrices. We now ask, with the PCA in mind, about the statistics of the largest eigenvalue λ_M. We expect that $\langle \lambda_M \rangle = N x_+(c)$ and from the joint pdf of the eigenvalues, we might even expect that the probability to find λ_M less than some threshold t has a large deviation form

$$\mathrm{Prob}\{\lambda_M \leq t\} \sim \mathrm{e}^{-\frac{1}{2}\beta N^2 \Phi_-\left(\frac{x_+ N - t}{N}\right)} \tag{45}$$

as long as t is of order N. This is just an educated guess at this stage. If this is true, then the probability that the largest eigenvalue is less than the mean is

$$\mathrm{Prob}\{\lambda_M \leq N/c\} \sim \mathrm{e}^{-N^2 \times \text{a function of } c} \tag{46}$$

As will appear below, it is actually rather remarkable that a method can be found to estimate the probability that the largest eigenvalue is less than a given threshold, because the eigenvalues form a set of strongly correlated random variables. Our starting point is the distribution of the rescaled eigenvalues $x_i = \frac{\lambda_i}{\beta N}$ which reads

$$\rho(x_1, \ldots, x_N) = \frac{1}{Z_0} \mathrm{e}^{-\beta E} \tag{47}$$

where, using $\hat{\rho}(x) = \frac{1}{N}\sum_i \delta(x - x_i)$

$$\frac{E}{N^2} = \int \mathrm{d}x \left[\frac{x}{2} - \frac{1}{2}(c^{-1} - 1)\ln x\right] \hat{\rho}(x) - \frac{1}{2}\int \mathrm{d}x \mathrm{d}x' \hat{\rho}(x) \ln|x - x'| \hat{\rho}(x') \tag{48}$$

The normalization is given by $Z_0 = \int \mathrm{d}x_1 \ldots \mathrm{d}x_N \mathrm{e}^{-\beta E}$. The integration interval of the x_i's runs from 0 to $+\infty$. But if we ask about the probability that the largest of the x_i's (denoted by x_{Max}) is actually less than a given threshold ζ, we just have to estimate

$$\mathrm{Prob}\{x_{\mathrm{Max}} \leq \zeta\} = \int_0^\zeta \mathrm{d}x_1 \ldots \int_0^\zeta \mathrm{d}x_N \rho(x_1, \ldots, x_N) \tag{49}$$

This can be rewritten in the form

$$\text{Prob}\{x_{\text{Max}} \leq \zeta\} = \frac{Z_1}{Z_0} \tag{50}$$

where

$$Z_1 = \int_0^\zeta \mathrm{d}x_1 \ldots \int_0^\zeta \mathrm{d}x_N \mathrm{e}^{-\beta E} \tag{51}$$

If we adopt the functional formulation, in which

$$Z_0 = \int \mathscr{D}\hat{\rho} \mathrm{e}^{N \int \mathrm{d}x \hat{\rho} \ln \hat{\rho}} \mathrm{e}^{-\beta N^2 F[\hat{\rho}]} \tag{52}$$

then, in quite a similar fashion, we can write that

$$Z_0 = \int \mathscr{D}\hat{\rho} \mathrm{e}^{N \int \mathrm{d}x \hat{\rho} \ln \hat{\rho}} \mathrm{e}^{-\beta N^2 F_1[\hat{\rho}]} \tag{53}$$

with

$$F_1[\rho] = \frac{1}{2} \int^\zeta \mathrm{d}x (x - (1/c - 1) \ln x) \hat{\rho}(x) - \frac{1}{2} \int^\zeta \mathrm{d}x \mathrm{d}x' \hat{\rho}(x) \hat{\rho}(x') \ln |x - x'| \tag{54}$$

which differs from F by the upper bound of the x integrals, which is now confined to $[0, \zeta]$. We know that the optimal profile (unconstrained) has support $[\zeta_-, \zeta_+]$, so that we expect that if $\zeta > \zeta_+$, $Z_1 = Z_0$, but for $\zeta_- < \zeta < \zeta_+$, we need to get back to the explicit computation of the saddle, that is the solution of $\frac{\delta F_1}{\delta \rho} = 0$. The only difference now is that the Tricomi integral equation has ζ as an uppoer bound that we impose. This will of course lead to a ζ-dependent profil $\rho_1(x)$, and the evaluation of $F_1[\rho_1]$ will give us a pure number that depends on ζ only (and on the parameter c). This confirms that

$$\text{Prob}\{x_{\text{Max}} \leq \zeta\} \sim \mathrm{e}^{-N^2 \times \text{a function of } \zeta \text{ and } c} \tag{55}$$

The explicit calculations can be read in the reference given at the begining of this subsection.

4 Exercises

4.1 Did you realize that...

- establishing the normalization factor in Eq. (1) is nontrivial. Prove that formula.
- establishing Eq. (13) is nontrivial. Prove that formula.

Solution: Let's proceed backwards and start from the formula:

$$
\begin{aligned}
\mathrm{Tr}(TK) &= \sum_{i,j} T_{ij}K_{ji} \\
&= \sum_i T_{ii}K_{ii} + \sum_{i<j}((T^R_{ij} + iT^I_{ij})(K^R_{ij} + iK^I_{ij}) + \sum_{j<i}((T^R_{ij} + iT^I_{ij})(K^R_{ij} + iK^I_{ij}) \\
&= \sum_i T_{ii}K_{ii} + \sum_{i<j}((T^R_{ij} + iT^I_{ij})(K^R_{ij} + iK^I_{ij}) + \sum_{i<j}((T^R_{ji} + iT^I_{ji})(K^R_{ji} + iK^I_{ji}) \\
&= \sum_i T_{ii}K_{ii} + \sum_{i<j}((T^R_{ij} + iT^I_{ij})(K^R_{ij} + iK^I_{ij}) + \sum_{i<j}((T^R_{ij} - iT^I_{ij})(K^R_{ij} - iK^I_{ij}) \\
&= \sum_i T_{ii}K_{ii} + 2\sum_{i<j}(T^R_{ij}K^R_{ij} + T^I_{ij}K^I_{ij})
\end{aligned}
\tag{56}
$$

and then we exponentiate and integrate. The diagonal elements bring in a factor $\prod_{i=1}^{N}[2\pi\delta(K_{ii})]$. The integration over the $N(N-1)/2$ real parts leads to $\prod_{i<j} 2\pi\delta(2K^R_{ij}) = \prod_{i<j}\left[\pi\delta(K^R_{ij})\right]$. A similar factor appears for the imaginary parts, and thus altogether one gets $(2\pi)^N \times \pi^{N(N-1)/2} \times \pi^{N(N-1)/2} \times \delta(K)$. This is exactly Eq. (13).

- establishing Eq. (20) is nontrivial. Prove that formula.

Solution: We want to prove that

$$
\int \mathrm{d}T \mathrm{e}^{\frac{i}{2}\mathrm{Tr}(TW)} \det(i\mathbf{1}_{N,N} - T)^{-M} \propto (\det W)^{M-N} \mathrm{e}^{-\frac{1}{2}\mathrm{Tr}\,W} \tag{57}
$$

This is done by shifting...

- establishing Eq. (28) is nontrivial. Prove that formula by discretizing space and working on a lattice.

4.2 The origin of the $|\lambda_i - \lambda_j|^\beta$ factor in the joint pdf of the eigenvalues

The elements M_{ij} of a real symmetric random matrix M of size $N \times N$ are distributed according to a probability density that can be expressed as a function of $\mathrm{Tr}M^2$ only ($p(\mathrm{Tr}M^2)$).

1. How many independent elements does the matrix M possess?

Solution: It has $\frac{N(N+1)}{2}$ independent elements (N on the diagonal, and $N(N-1)/2$ above the diagonal.

2. Let λ_i, $i = 1, \ldots, N$ the N eigenvalues of a given matrix M. What is the argument that allows to assert that M can be diagonalized? Express $\mathrm{Tr}M$ and $\mathrm{Tr}M^2$ in terms of the eigenvalues of M.

Solution: We have $\mathrm{Tr}M = \sum_i \lambda_i$ and $\mathrm{Tr}M^2 = \sum_i \lambda_i^2$ (because $M = R\varLambda R^{-1}$ and $M^2 = R^2\varLambda^2 R^{-1}$).The spectral theorem applies: M can be diagonalized in an orthonormal basis made of its eigenvectors ($R \in O(N)$).

3. The matrix M is written in the form

$$M = R \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} R^{-1} \tag{58}$$

What can be said about the matrix R? This is also a random matrix. Justify that it possesses $N(N-1)/2$ independent elements (or that a rotation of $O(N)$ is parametrized by means of $N(N-1)/2$ coefficients).

Solution: The matrix R is orthogonal, $R^T = R^{-1}$, which imposes a number of constraints upon it $N \times N$ elements. Let A such that $R = \mathrm{e}^{\varOmega}$, then we see that $\varOmega$ is antisymmetric which leaves us with $\frac{N(N-1)}{2}$ independent parameters. Taking the N eigenvalues into account we do find that $N + \frac{N(N-1}{2} = \frac{N(N+1)}{2}$.

4. Our interest goes to a quantity F that depends on M only through the eigenvalues of M. It is then interesting to perform a change of variables from the $\frac{N(N+1)}{2}$ independent elements M_{ij} to the N eigenvalues λ_i and the $N(N-1)/2$ independent elements of $\varOmega$ in question **??**. Let $J(\{\lambda_i\}, R)$ be the Jacobian of this transformation. Prove that if a symmetric matrix M is varied by $\mathrm{d}M$ then

$$\mathrm{d}M = R(\mathrm{d}\varOmega\varLambda - \varLambda\mathrm{d}\varOmega + \mathrm{d}\varLambda)R^T,\, \mathrm{d}\varOmega = R^T\mathrm{d}R \tag{59}$$

where $\varLambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$ is the matrix with the eigenvalues of M on the diaginal and R is matrix of the related eigenvectors.

5. Let $\mathrm{d}M' = \mathrm{d}\varOmega\varLambda - \varLambda\mathrm{d}\varOmega + \mathrm{d}\varLambda$. Prove that

$$\frac{\partial M'_{ij}}{\partial \lambda_k} = \delta_{ij}\delta_{ik},\ \frac{\partial M'_{ij}}{\partial \varOmega_{k\ell}} = \delta_{ik}\delta_{jl}(\lambda_j - \lambda_i) \tag{60}$$

6. How can we now conclude that the Jacobian of the transformation is of the form $\prod_{i<j} |\lambda_i - \lambda_j|$ up to N-dependent constant?

   Solution: A Jacobian is by definition a determinant, that of an $\frac{N(N+1)}{2} \times \frac{N(N+1)}{2}$ matrix. The matrix is diagonal, the formula follows from ultiplying all diagonal elements.

7. What is the $N \times N$ matrix whose determinant is $\prod_{i<j}(\lambda_i - \lambda_j)$?

   Solution: This is a Vandermonde matrix V with elements $V_{ij} = \lambda_i^{j-1}$, $i, j = 1, \ldots, N$. The determinant is a polynomial of degree $\frac{N(N-1)}{2}$ (and of degree $N-1$ in each λ_i) which vanishes when two eigenvalues are identical (then V has two identical lines) and the leading term is known. This completely fixes the form of the determinant.

4.3 Largest eigenvalue in the GOE

What can you say about its statistics?

Solution: It scales as $\sqrt{2N} + N^{-1/6}x$ where $x \sim O(1)$.