WHAT IS...

a Random Matrix?

Persi Diaconis

When I became a professor at Harvard, David Kazhdan was running a "basic notions" seminar—things every graduate student (and perhaps also faculty member) should know. He asked me to give a talk on "what is a random variable." Since a random matrix is a random variable taking values in the space of matrices, perhaps it's good to start with random variables.

I was surprised by Kazhdan's request since "everybody knows" that a random variable is just a measurable function

$$X(\omega) \text{ from } \Omega \text{ to } \mathcal{X}.$$

He answered "yes, but that's not what it means to people working in probability" and of course he was right. Let us consider the phrase

Pick a random matrix from Haar measure
on the orthogonal group O_n

Here O_n is the group of $n \times n$ real matrices X with $XX^T = id$. Most of us learn that O_n has an invariant probability measure μ, that is, a measure on the Borel sets A of O_n such that for every set A and matrix M

$$\mu(A) = \mu(MA), \ \mu(O_n) = 1.$$

Let me tell you how to "pick X from μ." To begin with, you will need a sample Y_{ij} of picks from the standard normal density (bell-shaped curve). This is the measure on the real line with density $\frac{e^{-x^2/2}}{\sqrt{2\pi}}$. Even if you don't know what it means to "pick Y_{ij} from the normal density," your computer knows. You can just push a button and get a stream of independent normal picks.

Now things are easy. Fill up an empty $n \times n$ array with Y_{ij}, $1 \le i, j \le n$. Turn this into an orthogonal matrix by applying the Gram-Schmidt algorithm; make the first row have norm one, take the first row out of the second row and then make this have norm one, and so on. The resulting matrix X is random (because it was based on the random Y_{ij}) and orthogonal (because we forced it to be). Using the orthogonal invariance of the normal distribution it is not hard to prove that X has the invariant Haar measure

$$\text{probability}(X \in A) = \mu(A).$$

Let us now translate the algorithmic description of a random orthogonal matrix into random variable language. Let $\Omega = \mathbb{R}^{n^2}$. Let $\mathcal{X}$ be the orthogonal group. The Gram-Schmidt algorithm gives a map $X(\omega)$ from almost all of Ω onto $\mathcal{X}$. This $X(\omega)$ is our random variable. To prescribe its probability distribution, put product measure of $\frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$ on $\mathbb{R}^{n^2}$. The push forward of this measure under the map X is Haar measure μ.

Often, one is interested in the eigenvalues of the matrix. For orthogonal matrices, these are n points on the unit circle. Figure 1a shows the eigenvalues of a random 100×100 orthogonal matrix. While there is some local variation, the eigenvalues are very neatly distributed. For contrast, Figure 1b shows 100 points put down independently at random on the unit circle. There are holes and clusters that do not appear in Figure 1a. For details, applications and a lot of theory supplementing these observations, see Diaconis (2003).

So far, I have answered the question "what is a random orthogonal matrix?" For a random unitary matrix replace the normal distribution on $\mathbb{R}$ with the normal distribution on $\mathbf{C}$. This has density $\frac{e^{-|z|^2}}{\pi} dz$. We choose a random complex normal variable Z on a computer by choosing real independent normals Y_1 and Y_2 and setting

Persi Diaconis is the Mary V. Sunseri Professor of Statistics and Mathematics at Stanford University.

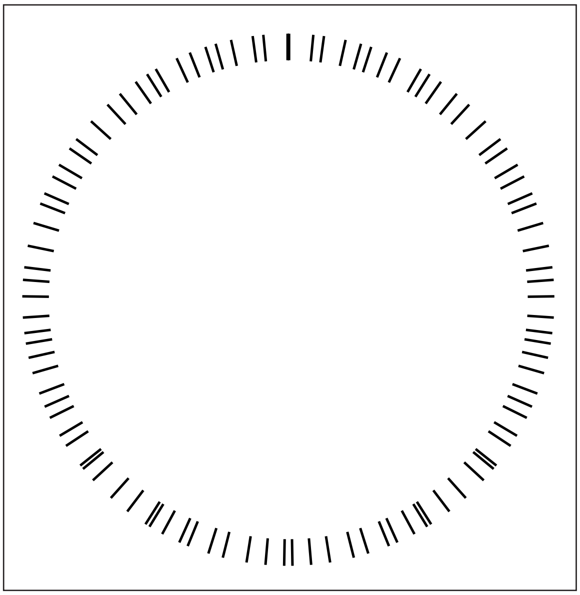

Figure 1a. Eigenvalues of a random orthogonal matrix.

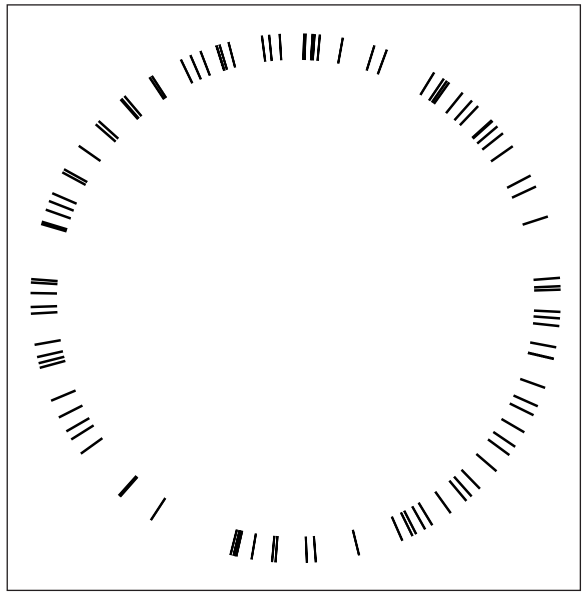

Figure 1b. Random points on the circle.

$Z = \frac{1}{2}Y_1 + i\frac{1}{2}Y_2$. For a random symplectic matrix, use the quaternions and $Z = \frac{1}{4}(Y_1 + iY_2 + jY_3 + kY_4)$. Random matrices in orthogonal, unitary, and symplectic groups are called the classical circular ensembles in the physics literature.

There are also noncompact ensembles; to choose a random Hermitian matrix, fill out an $n \times n$ array by putting picks from the standard complex normal above the diagonal, picks from the standard real normal on the diagonal, and finally filling below the diagonal by using complex conjugates of what is above the diagonal. This is called GUE (the Gaussian unitary ensemble) in the physics literature because the random matrices have distribution invariant under multiplication by the unitary group. There are other useful ensembles considered in the classical book by Mehta (2004). One of the interesting claims argued there is that only three universal families (orthogonal, unitary, and symplectic) need be considered; many large n problems have answers the same as for these families, no matter what probability distribution governs the matrices involved.

Historically, random matrix theory was started by statisticians studying correlations between different features of a population (height, weight, income...). This led to correlation matrices with (i, j) entry the correlation between the ith and jth features. If the data was based on a random sample from a larger population, these correlation matrices are random; the study of how the eigenvalues of such samples fluctuate was one of the first great accomplishments of random matrix theory. These values and the associated eigenvectors are a mainstay of a topic called "principle components analysis". This seeks low-dimensional descriptions of high-dimensional data; it is widely used across applied areas from psychology to oceanography and is a crucial ingredient of search engines such as Google. See Diaconis (2003) for more details.

Physicists began to study random matrix theory in the 1950s as a useful description of energy differences in things like slow neutron scattering. This has grown into an enormous literature which has been developed to study new materials (quantum dots) and parts of string theory. There have also been wonderful applications of random matrix theory in combinatorics and number theory. One can find the literature on all of these topics by browsing in Forrester et al (2003).

Returning finally to random matrices as mathematical objects, the reader will see that we have been treating them as "real" rather than as abstract mathematical objects. From a matrix one passes to the eigenvalues and then perhaps to a spectrum renormalized to have average spacing one and then to the histogram of spacings. All of this is mechanically translatable to the language of measurable functions. However, this is a bit like working in assembly language instead of just naturally programming your Mac. Random matrix theory has evolved as the high order descriptive language of this rich body of results.

References

P. Diaconis, Patterns in Eigenvalues: The 70th Josiah Willard Gibbs Lecture, *Bull. Amer. Math. Soc.* **40**, 155–178 (2003).

P. Forrester, N. Snaith, and V. Verbaarschot, Introduction Review to Special Issue on Random Matrix Theory, *Jour. Physics A: Mathematical and General* **36**, R1–R10 (2003).

M. Mehta, *Random Matrices*, 3rd Edition, Elsevier, Amsterdam. (2004).