The road from Kac's matrix to Kac's random polynomials

Alan Edelman*   Eric Kostlan†

Abstract

This paper tells the story of a matrix and a problem both of which are associated with the name Mark Kac, though most likely he never made the connection. This matrix appears as the Clement matrix in Higham's toolbox, and in the Ehrenfest urn model of diffusion. Our immediate interest in this matrix arises in the study of roots of random polynomials.

1 A<—>B

This report may be thought of as a path with two endpoints, both of which may be associated with the name Mark Kac. No on second thought, the focus is not so much on the path, but rather the two endpoints themselves as there are many paths between the two points, some more scenic than others. We have no reason to suspect that Kac ever took the path from **A** to **B** as his itinerary took him to many other destinations.

Endpoint "A" is the so-called Kac matrix, a simple matrix with many hidden treasures. Endpoint "B" is a Kac problem on the average number of real roots of a random algebraic equation. The matrix ("A") appeared in 1946 in a lecture on Brownian motion [12]. The random polynomial question ("B") first appeared in a 1943 paper [11].

2 The Kac Matrix

The $n+1$ by $n+1$ Kac matrix is defined as the tridiagonal matrix

$$\mathrm{Kac}_n \equiv \begin{pmatrix} 0 & n & & & & \\ 1 & 0 & n-1 & & & \\ & 2 & 0 & n-2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & n-1 & 0 & 1 \\ & & & & n & 0 \end{pmatrix}.$$

This matrix was dubbed "A Matrix of Mark Kac" by Taussky and Todd [18] who point out that this matrix was studied by Sylvester, Schrödinger, and many others. It also has the name "Clement matrix" in Higham's *Test Matrix Toolbox for Matlab* [10] because of Clement's [2] proposed use of this matrix as a test matrix. It is the matrix that describes a random walk on a hypercube as well as the Ehrenfest urn model of diffusion [3, 4].

*Department of Mathematics Room 2-380, Massachusetts Institute of Technology, Cambridge, MA 02139, `edelman@math.mit.edu`.

†Arts and Sciences, Kapi'olani Community College, 4303 Diamond Head Road, Honolulu, HI 96816, `kostlan@math.hawaii.edu`.

The first surprise that this matrix has in store for us is that the eigenvalues are the integers $-n, -n+2, -n+4, \ldots, n-2, n$. Several proofs of this fact may be found in [18] along with interesting historical remarks such as that Schrödinger could not find a proof that these eigenvalues are correct. Numerically the eigenvalues can be difficult in a manner that reminds us of the Wilkinson matrix. Though the eigenvalues are integers in exact arithmetic, Matlab computes some non-real eigenvalues when `eig(clement(120))` is executed.

One of the proofs in [18] is denoted "mild trickery by Kac" in that Kac makes clever use of generating functions to obtain the eigenvalues of eigenvectors. We like to think that if Kac's proof was mild trickery, the following new proof might be considered major trickery in that the reader would be unable to guess how we conjured up this proof.

THEOREM 2.1. *The eigenvalues of* Kac_n *are the integers* $2k-n$ *for* $k = 0, 1, \ldots, n$.

Proof. Define

$$f_k(x) \equiv \sinh^k(x)\cosh^{n-k}(x), \quad k = 0, \ldots, n,$$

$$g_k(x) \equiv (\sinh(x) + \cosh(x))^k (\sinh(x) - \cosh(x))^{n-k}, \quad k = 0, \ldots, n$$

If V is the vector space of functions with basis $\{f_k(x)\}$, then the $g_k(x)$ are clearly in this vector space. Also, $\frac{d}{dx} f_k(x) = k f_{k-1}(x) + (n-k) f_{k+1}(x)$, so that the Kac matrix is the representation of the operator d/dx in V. We actually wrote $g_k(x)$ in a more complicated way than we needed to so that we could emphasize that $g_k(x) \in V$. Actually, $g_k(x) = \exp((2k-n)x)$ is an eigenfunction of d/dx with eigenvalue $2k-n$ for $k = 0, \ldots, n$. The eigenvector is obtained by expanding out the $g_k(x)$ in terms of the $f_k(x)$. □

3 The Kac-Matrix as Polynomial "Rotators"

Perhaps a good magician never reveals his secrets, but we are mathematicians, and we can not resist. Our secret is a certain understanding of nth degree polynomials in one variable. As usual, we will identify the set of polynomials of the form $p(t) = a_0 + a_1 t + \cdots + a_n t^n$, with the vector space $\Re^{n+1}$. Actually, it will be more convenient to "homogenize" the polynomial so as to consider polynomials of the form $p(x, y) = a_0 x^n + a_1 x^{n-1} y + \cdots + a_n y^n$ which we will not distinguish from the vector $(a_0, a_1, \ldots, a_n)^T$.

For each $\theta \in [0, 2\pi)$, we define the "rotated" polynomial

$$p(x, y; \theta) \equiv p(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta).$$

With all the sines and cosines around, it would be easy to lose sight of the fact that for the coefficients of the polynomial $p(x, y; \theta)$ are linear functions of the coefficients a_i of $p(x, y)$ even if they are complicated trigonometric functions in θ. Since we are not distinguishing a polynomial from the vector of its coefficients, we may say that there is an $n+1 \times n+1$ matrix L_θ for which

$$p(x, y; \theta) = L_\theta p(x, y).$$

We will demonstrate that the matrices L_θ are matrix exponentials of scalar multiples of the Kac-like matrix

$$\mathrm{AntiKac}_n \equiv \begin{pmatrix} 0 & n & & & & \\ -1 & 0 & n-1 & & & \\ & -2 & 0 & n-2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -(n-1) & 0 & 1 \\ & & & & -n & 0 \end{pmatrix}.$$

Indeed by differentiating such terms as $a_k(x\cos\theta + y\sin\theta)^{n-k}(-x\sin\theta + y\cos\theta)^k$ with respect to θ, we see that

$$\frac{d}{d\theta}L_\theta = \text{AntiKac}_n L_\theta,$$

which has the solution $L_\theta = \exp(\theta\text{AntiKac}_n)$.

If we let $D_n = \text{diag}(1, i, i^2, \ldots, i^n)$, then $D_n^{-1}\text{Kac}_n D_n = i\text{AntiKac}_n$. So the eigenvalues of AntiKac_n are the integers $-n, -n+2, \ldots, n-2, n$ multiplied by i. The reader may wish to return to the definition of the rotated polynomials to see why the eigenvalues "had to" be integers times i.

Of course the matrices Kac_n and AntiKac_n are really the same. The difference is just the difference between sinh and sin, it all depends on which way you are facing in the complex plane.

4 The symmetrized Kac matrix and Random Polynomials

The matrix Kac_n (or AntiKac_n) may be symmetrized (anti-symmetrized) by a diagonal matrix containing square-roots of the binomial coefficients: $B_n = \text{diag}(\{\binom{n}{k}^{1/2}\}_{k=0}^n)$ The symmetrized or version of the matrix contains the numbers $\sqrt{k(n+1-k)}$ on the super and sub-diagonals; the anti-symmetrized is the same except that the subdiagonal entries have minus signs.

Let us to go back to our definition in Section 3 and say that we will now identify $\sum a_k\binom{n}{k}^{1/2}x^{n-k}y^k$ with the vector $(a_0, a_1, \ldots, a_n)^T$. This is a scaling of the coordinate axes of our $n+1$ dimensional space. If we were to follow our definitions, the matrices L_θ are now exponentials of anti-symmetric matrices, i.e. they are orthogonal. In other words, we may compute the coefficients of $p(x, y; \theta)$ by applying an orthogonal matrix to the coefficients of $p(x, y)$. In this coordinate system, rotating the homogeneous arguments of p induces a rotation of the coefficient vector of p! As θ sweeps through $[0, 2\pi)$, the coefficient vector of the polynomial sweeps out a path in $\Re^{n+1}$ that is confined to a sphere centered around the origin. In general, this path will be non-planar. One interesting degenerate case is the polynomial $(x^2+y^2)^{n/2}$ defined for even n. Rotating this polynomial does not move it; this polynomial is the eigenvector corresponding to the eigenvalue 0. Other degenerate cases that do lead to planar paths (circles centered about the origin), may be obtained from the eigenvectors corresponding to $\pm ki$.

What if $p(x, y) = \sum a_k x^{n-k}y^k$ is a random polynomial with coefficients a_k taken from independent and identically distributed standard normal distributions? It is well known that the distribution of a vector of independent standard normals is spherically symmetric. Since L_θ is orthogonal, we see that the distribution of the vector of coefficients of $p(x, y; \theta)$ is the same as that of $p(x, y)$. The probability distribution of the coefficients of our rotated polynomials is the same as that of our original polynomials!

Therefore, the probability distribution of the roots of the polynomials is invariant under rotation. Remembering that if (x, y) is a root of the homogenized polynomial, then $t = y/x$ is a root of the unhomogenized polynomial, we see that $\arctan(t)$ is uniformly distributed on $[-\pi/2, \pi/2)$.

We therefore conclude THEOREM 4.1. *If $p(t) = \sum a_k\binom{n}{k}^{1/2}t^k$ is a random polynomial with normally distributed coefficients, then the distribution of the real roots of $p(t) = 0$ has the Cauchy distribution, i.e.* $\arctan(t)$ *is uniformly distributed on* $[-\pi/2, \pi/2)$.

5 A Curve Length Counts How Many Random Roots are Real

Let $f_k(t), k = 0, 1, \ldots, n$ be a collection of rectifiable functions

$$v(t) = \begin{pmatrix} f_0(t) \\ f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix},$$

and let $\gamma(t) = v(t)/\|v(t)\|$ so $\gamma(t)$ is on the unit sphere. We showed in [8] that the expected number of real roots t in the interval $[a, b]$ to the random equation $\sum_{k=0}^{n} a_k f_k(t) = 0$, a_k independent standard normals, is $1/\pi$ times the length of the curve $\gamma(t)$ that is swept out on the sphere as t runs through $[a, b]$. A proof of this statement using mostly precalculus level mathematics may be found in the reference.

Kac [11, 12] considered the first question that might come to a reader's mind: $f_k(t) = t^k$. As $n \to \infty$ the expected number of real roots is asymptotic to $\frac{2}{\pi}\log n + 0.6257350972\ldots + \frac{2}{n\pi} + O(1/n^2)$ as $n \to \infty$. (Kac knew the leading behavior $\frac{2}{\pi}\log n$; his derivation was algebraic, not geometric.)

Here, we wish to focus on a random polynomial question introduced by Kostlan [14] that is more closely connected to the Kac matrix: $f_k(t) = \binom{n}{k}^{1/2} t^k$. We may build the vector $v(t)$, and then normalize to the unit sphere. Letting $t = \tan\theta$ simplifies the answer which is

$$\gamma(\tan(\theta)) = \begin{pmatrix} \binom{n}{0}^{1/2} \cos^n\theta \\ \binom{n}{1}^{1/2} \cos^{n-1}\theta \sin\theta \\ \binom{n}{2}^{1/2} \cos^{n-2}\theta \sin^2\theta \\ \vdots \\ \binom{n}{n}^{1/2} \sin^n\theta \end{pmatrix},$$

i.e. $\gamma_k(\theta) = \binom{n}{k}^{1/2} \cos^{n-k}\theta \sin^k\theta$, where the dimension index k runs from 0 to n. The binomial expansion of $(\sin^2\theta + \cos^2\theta)^n = 1$ checks that our curve lives on the unit sphere.

If we differentiate, $\gamma(\tan(\theta))$ with respect to θ the anti-symmetrized AntiKac_n matrix appears. The curve on the sphere traced out by $\gamma(\theta)$ is the same curve traced out by $p(x, y; \theta)$, when $p(x, y) = x^n$. Though θ varies, the velocity vector always has the same length as the first column of the anti-symmetrized AntiKac_n matrix, which is $\sqrt{n}$. We may conclude. THEOREM 5.1. *If $p(t) = \sum_{k=0}^{n} a_k \binom{n}{k}^{1/2} t^k$ is a random polynomial with normally distributed coefficients, then the expected number of real roots to the equation $p(t) = 0$ is exactly $\sqrt{n}$.*

Proof. As θ runs through $[-\pi/2, \pi/2)$, we trace out the curve $\gamma(\tan(\theta))$ of length $\pi\sqrt{n}$ because the speed of the curve at every point is $\sqrt{n}$. Dividing the result by π yields the result. □

In conclusion, we changed Kac's question a little by asking for the roots of the random equation $0 = \sum k = 0^n a_k \binom{n}{k}^{1/2} t^k$, and we found that a small variation on Kac's matrix may be found everywhere in the analysis.

6 Acknowledgments

We would like to thank Beresford Parlett for introducing us to the Kac matrix on a staircase in Evans hall at UC Berkeley. Without this fortuitous discussion we would never have known

that the matrix we were studying in the context of Kac's problem also was named for Kac.

References

[1] A.T. Bharucha-Reid and M. Sambandham, *Random Polynomials*, Academic Press, New York, 1986.

[2] P.A.Clement, A class of triple-diagonal matrices for test purposes, *SIAM Review, 1* (1959), 50–52.

[3] P. Diaconis, *Group Representations in Probability and Statistics*, Institute of Mathematical Statistics, Hayward, CA, 1988.

[4] P.Diaconis, R.L.Graham, and J.A.Morrison, Asymptotic analysis of a random walk on a hypercube with many dimensions, *Random Structures and Algorithms, 1* (1990), 51–72.

[5] A. Edelman, Eigenvalues and condition numbers of random matrices, *SIAM J. Matrix Anal. Appl. 9* (1988), 543–560.

[6] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, PhD thesis, Department of Mathematics, MIT, 1989.

[7] A. Edelman, E. Kostlan, and M. Shub, How many eigenvalues of a random matrix are real?, *J. Amer. Math. Soc. 7* (1994), 247–267.

[8] A. Edelman and E. Kostlan, How many zeros of a random polynomial are real?, preprint.

[9] A. Edelman, The circular law and the probability that a random matrix has k real eigenvalues, preprint.

[10] N. Higham, The Test Matrix Toolbox for Matlab, Numerical Analysis Report No. 237, University of Manchester, England, December 1993.

[11] M. Kac, On the average number of real roots of a random algebraic equation, *Bull. Am. Math. Soc. 49* (1943), 314–320 and 938.

[12] M. Kac, On the average number of real roots of a random algebraic equation (II), *Proc. London Math. Soc. 50* (1949), 390–408.

[13] M. Kac, Random walk and the theory of Brownian motion, *Bull. Amer. Math. Soc. 54* (1947), 369–391.

[14] E. Kostlan, On the distribution of roots of random polynomials, Chapter 38 (pp. 419-431) of *From Topology to Computation: Proceedings of the Smalefest* edited by M.W. Hirsch, J.E. Marsden and M. Shub, Springer-Verlag, New York, 1993.

[15] A.M. Odlyzko and B. Poonen, Zeros of Polynomials with 0,1 Coefficients, *J. L'Enseign. Math.*, to appear.

[16] L.A. Santaló, *Integral Geometry and Geometric Probability,* Volume 1 of *Encyclopedia of Mathematics and Its Applications,* Addison-Wesley, Reading, 1976.

[17] M. Shub and S. Smale, Complexity of Bezout's Theorem II: Volumes and Probabilities, in *Computational Algebraic Geometry*, F. Eyssette and A. Galligo, eds, Progress in Mathematics, v 109, Birkhauser, 1993, 267–285.

[18] O. Taussky and J. Todd, Another look at a matrix of Mark Kac, *Linear Algebra Appl.* **150**, (1991), 341–360.