User-Friendly Tools for Random Matrices

Joel A. Tropp

Computing + Mathematical Sciences
California Institute of Technology
jtropp@cms.caltech.edu

Research supported by ONR, AFOSR, NSF, DARPA, Sloan, and Moore.

Download the Notes:

tinyurl.com/bocrqhe

[URL] `http://users.cms.caltech.edu/~jtropp/notes/Tro12-User-Friendly-Tools-NIPS.pdf`

Random Matrices
in the Mist

Random Matrices in Statistics

- Covariance estimation for the multivariate normal distribution

John Wishart

3. *Multi-variate Distribution. Use of Quadratic co-ordinates.*

A comparison of equation (8) with the corresponding results (1) and (2) for uni-variate and bi-variate sampling, respectively, indicates the form the general result may be expected to take. In fact, we have for the simultaneous distribution in random samples of the n variances (squared standard deviations) and the $\frac{n(n-1)}{2}$ product moment coefficients the following expression:

$$dp = \frac{\begin{vmatrix} A_{11} & A_{12} \dots & A_{1n} \\ A_{21} & A_{22} \dots & A_{2n} \\ \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} \dots & A_{nn} \end{vmatrix}^{\frac{N-1}{2}}}{(\sqrt{\pi})^{\frac{1}{2}n(n-1)} \Gamma\left(\frac{N-1}{2}\right)\Gamma\left(\frac{N-2}{2}\right)\dots\Gamma\left(\frac{N-n}{2}\right)}$$

$$\times e^{-A_{11}a_{11} - A_{22}a_{22} - \dots - A_{nn}a_{nn} - 2A_{12}a_{12} - 2A_{13}a_{13} - \dots - 2A_{n-1\,n}a_{n-1\,n}}$$

$$\times \begin{vmatrix} a_{11} & a_{12} \dots & a_{1n} \\ a_{21} & a_{22} \dots & a_{2n} \\ \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} \dots & a_{nn} \end{vmatrix}^{\frac{N-n-2}{2}} da_{11} da_{12} \dots da_{nn} \quad \dots\dots\dots\dots\dots\dots(9),$$

where $a_{pq} = s_p s_q r_{pq}$, and $A_{pq} = \frac{N}{2\sigma_p\sigma_q} \cdot \frac{\Delta_{pq}}{\Delta}$, Δ being the determinant

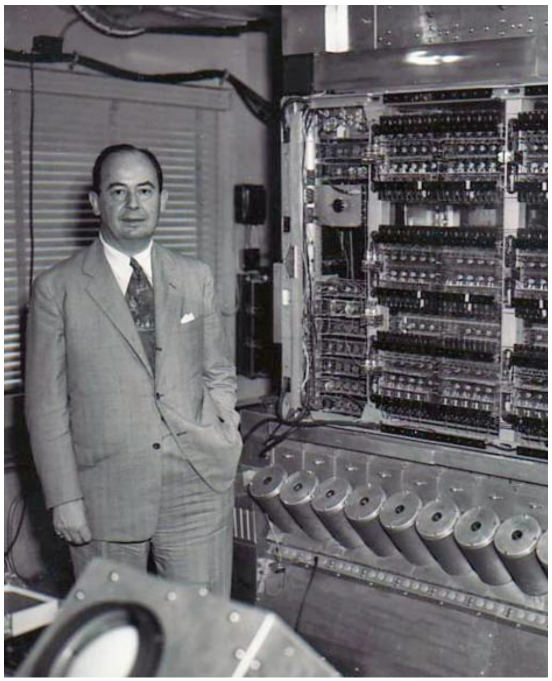
$$|\rho_{pq}|,\ p,\ q = 1, 2, 3, \dots n,$$

and Δ_{pq} the minor of ρ_{pq} in Δ.

[Refs] Wishart, *Biometrika* 1928. Photo from `apprendre-math.info`.

Random Matrices in Numerical Linear Algebra

❧ Model for floating-point errors in LU decomposition



John von Neumann

now combining (8.6) and (8.7) we obtain our desired result:

$$\text{(8.8)}\quad \text{Prob}\,(\lambda > 2\sigma^2 rn) < \frac{(rn)^{n-1/2}e^{-rn}\pi^{1/2}e^n\cdot 2^{n-2}}{\pi n^{n-1}(r-1)n} = \left(\frac{2r}{e^{r-1}}\right)^n \times \frac{1}{4(r-1)(r\pi n)^{1/2}}\,.$$

We sum up in the following theorem:

(8.9) The probability that the upper bound $|A|$ of the matrix A of (8.1) exceeds $2.72\sigma n^{1/2}$ is less than $.027\times 2^{-n}n^{-1/2}$, that is, with probability greater than 99% the upper bound of A is less than $2.72\sigma n^{1/2}$ for $n=2, 3, \cdots$.

This follows at once by taking $r=3.70$.

[Refs] von Neumann and Goldstine, *Bull. AMS* 1947 and *Proc. AMS* 1951. Photo ©IAS Archive.

Random Matrices in Nuclear Physics

❧ Model for the Hamiltonian of a heavy atom in a slow nuclear reaction

Eugene Wigner

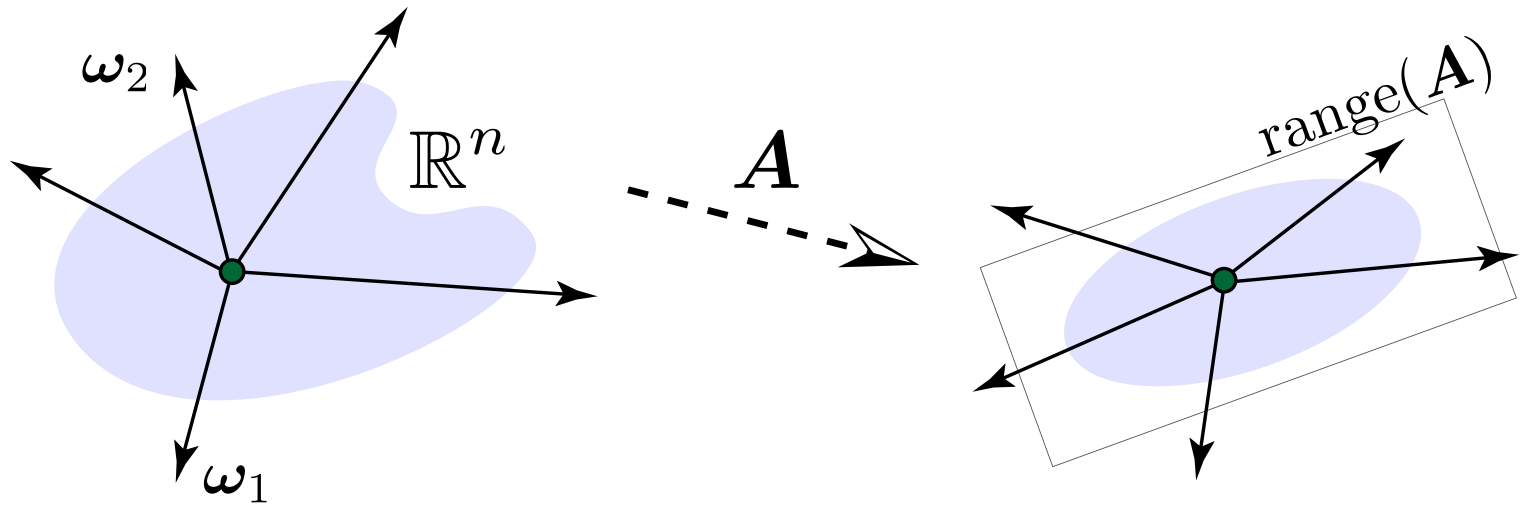
Random sign symmetric matrix

The matrices to be considered are $2N + 1$ dimensional real symmetric matrices; N is a very large number. The diagonal elements of these matrices are zero, the non diagonal elements $v_{ik} = v_{ki} = \pm v$ have all the same absolute value but random signs. There are $\mathfrak{N} = 2^{N(2N+1)}$ such matrices. We shall calculate, after an introductory remark, the averages of $(H^\nu)_{00}$ and hence the strength function $S'(x) = \sigma(x)$. This has, in the present case, a second interpretation: it also gives the density of the characteristic values of these matrices. This will be shown first.

[Refs] Wigner, *Ann. Math* 1955. Photo from Nobel Foundation.

Modern Applications

Randomized Linear Algebra



Input: An $m \times n$ matrix $\boldsymbol{A}$, a target rank k, an oversampling parameter p

Output: An $m \times (k+p)$ matrix $\boldsymbol{Q}$ with orthonormal columns

1. Draw an $n \times (k+p)$ random matrix $\boldsymbol{\Omega}$
2. Form the matrix product $\boldsymbol{Y} = \boldsymbol{A\Omega}$
3. Construct an orthonormal basis $\boldsymbol{Q}$ for the range of $\boldsymbol{Y}$

[Ref] Halko–Martinsson–T, *SIAM Rev.* 2011.

Other Algorithmic Applications

- **Sparsification.** Accelerate spectral calculation by randomly zeroing entries in a matrix.

- **Subsampling.** Accelerate construction of kernels by randomly subsampling data.

- **Dimension Reduction.** Accelerate nearest neighbor calculations by random projection to a lower dimension.

- **Relaxation & Rounding.** Approximate solution of maximization problems with matrix variables.

[Refs] Achlioptas–McSherry 2001 and 2007, Spielman–Teng 2004; Williams–Seeger 2001, Drineas–Mahoney 2006, Gittens 2011; Indyk–Motwani 1998, Ailon–Chazelle 2006; Nemirovski 2007, So 2009...

Random Matrices as Models

- **High-Dimensional Data Analysis.** Random matrices are used to model multivariate data.

- **Wireless Communications.** Random matrices serve as models for wireless channels.

- **Demixing Signals.** Random model for incoherence when separating two structured signals.

[Refs] Bühlmann and van de Geer 2011, Koltchinskii 2011; Tulino–Verdú 2004; McCoy–T 2011.

Theoretical Applications

- **Algorithms.** Smoothed analysis of Gaussian elimination.
- **Combinatorics.** Random constructions of expander graphs.
- **High-Dimensional Geometry.** Structure of random slices of convex bodies.
- **Quantum Information Theory.** (Counter)examples to conjectures about quantum channel capacity.

[Refs] Sankar–Spielman–Teng 2006; Pinsker 1973; Gordon 1985; Hayden–Winter 2008, Hastings 2009.

Random Matrices:

My Way

The Conventional Wisdom

“Random Matrices are Tough!”

[Refs] `youtube.com/watch?v=NO0cvqT1tAE`, most monographs on RMT.

Principle A

"But...

In many applications, a random matrix can be decomposed as a sum of independent random matrices:

$$\boldsymbol{Z} = \sum_{k=1}^{n} \boldsymbol{S}_k$$

Principle B

and

There are exponential concentration inequalities for the spectral norm of a sum of independent random matrices:

$$\mathbb{P}\left\{\|\boldsymbol{Z}\| \geq t\right\} \leq \exp\left(\quad \cdots \quad\right)$$

!!!”

The Vision

- **Challenge:** Random matrices are tough!
- **Approach:**
  - Write the random matrix as a sum of independent random matrices
  - Apply "packaged" concentration inequalities
- **Tradeoff:**

  [+] Wide range of applicability
  [+] Simplicity
  [−] Potential loss in accuracy

To learn more...

E-mail: `jtropp@cms.caltech.edu`
Web: `http://users.cms.caltech.edu/~jtropp`

Some papers:

- "User-friendly tail bounds for sums of random matrices," *FOCM*, 2011.
- "User-friendly tail bounds for matrix martingales." Caltech ACM Report 2011-01.
- "Freedman's inequality for matrix martingales," *ECP*, 2011.
- "A comparison principle for functions of a uniformly random subspace," *PTRF*, 2011.
- "From the joint convexity of relative entropy to a concavity theorem of Lieb," *PAMS*, 2012.
- "Improved analysis of the subsampled randomized Hadamard transform," *AADA*, 2011.
- "Tail bounds for all eigenvalues of a sum of random matrices" with A. Gittens. Submitted 2011.
- "The masked sample covariance estimator" with R. Chen and A. Gittens. *I&I*, 2012.
- "Matrix concentration inequalities..." with L. Mackey et al.. Submitted 2012.
- "User-Friendly Tools for Random Matrices: An Introduction." 2012.
- "Deriving matrix concentration inequalities..." with D. Paulin and L. Mackey. Submitted 2013.
- "Subadditivity of matrix φ-entropy..." with R. Chen. Submitted 2013.