Random points in the unit ball of ℓ_p^n

Guillaume Aubrun

Abstract. We show that two limit results from random matrix theory, due to Marčenko–Pastur and Bai–Yin, are also valid for matrices with independent rows (as opposed to independent entries in the classical theory), when rows are uniformly distributed on the unit ball of ℓ_p^n, under proper normalization.

Mathematics Subject Classification (2000). Primary 15A52, 52A21.

Keywords. ℓ_p^n spaces, random matrices, random vectors.

Let us start with the following classical results from Random Matrix Theory. Let Z be a random variable such that

$$\mathbf{E}Z = 0 \quad \text{and} \quad \mathbf{E}Z^2 = 1. \tag{1}$$

Consider an infinite array (Z_{ij}) of i.i.d. copies of Z. For each couple (n, N) of integers, let $G_{n,N}$ be the $N \times n$ random matrix

$$G_{n,N} = \left(\frac{1}{\sqrt{N}} Z_{ij} \right)_{1 \leqslant i \leqslant N, 1 \leqslant j \leqslant n}. \tag{2}$$

We consider also the matrix

$$A_{n,N} = G_{n,N}^\dagger G_{n,N}.$$

We may drop subscripts and write simply G and A. The matrix A is sometimes called a *sample covariance matrix*. Let $(\lambda_i(A))$ be the eigenvalues of A, arranged in decreasing order. We write $\lambda_{\max}(A)$ for $\lambda_1(A)$ and $\lambda_{\min}(A)$ for $\lambda_n(A)$. The *spectral measure* of A is the probability measure on $\mathbf{R}$ defined as

$$\mu_A = \frac{1}{n} \sum_i \delta_{\lambda_i(A)}.$$

In other words, $\mu_A(B)$ is the proportion of eigenvalues of A that fall in a Borel set $B \subset \mathbf{R}$. The following theorems describe the limit behaviour of the spectrum of such large-dimensional matrices, in both global and local regime.

Research was supported in part by the European Network PHD, FP6 Marie Curie Actions, MCRN-511953 and was done in part while the author was visiting the University of Athens.

Theorem (Marčenko–Pastur [6]**).** *Let (n,N) be a sequence of sizes tending to infinity in such a way that the ratio n/N has a limit $\beta \in]0,1[$. Then, almost surely, the sequence of (empirical) spectral measures $(\mu_{A_{n,N}})$ converges weakly to the deterministic measure $\mu_{(\beta)}$ supported on the segment*

$$[\lambda_-(\beta), \lambda_+(\beta)] := [(1-\sqrt{\beta})^2, (1+\sqrt{\beta})^2]$$

and with density

$$\frac{d\mu_{(\beta)}}{dx} = \frac{1}{2\pi\beta x}\sqrt{(x-\lambda_-(\beta))(\lambda_+(\beta)-x)}.$$

Theorem (Bai–Yin [2]**).** *Assume moreover that $\mathbf{E}Z^4 < \infty$. Let (n,N) be a sequence of sizes tending to infinity in such a way that the ratio n/N has a limit $\beta \in]0,1[$. Let $A_{n,N} = G^\dagger_{n,N} G_{n,N}$ with $G_{n,N}$ defined as (2). Then almost surely, we have*

$$\lim_{n,N\to\infty} \lambda_{\max}(A_{n,N}) = \lambda_+(\beta) = (1+\sqrt{\beta})^2,$$

$$\lim_{n,N\to\infty} \lambda_{\min}(A_{n,N}) = \lambda_-(\beta) = (1-\sqrt{\beta})^2.$$

In some cases it is natural to consider a more general model of random matrices, and to weaken the hypothesis “independence of entries” to “independence of rows”. A random vector X in $\mathbf{R}^n$ is said to be *isotropic* if for every direction $\theta \in S^{n-1}$,

$$\mathbf{E}\langle X, \theta\rangle = 0 \ \text{ and } \ \mathbf{E}\langle X, \theta\rangle^2 = 1. \tag{3}$$

It is actually enough to check (3) for θ being a vector of the canonical basis. Condition (3) is the analogue of condition (1). Note also that any random vector (unless it belongs almost surely to an affine hylperplane) has an affine image which is isotropic. Also, (3) can be rephrased as $\mathbf{E}X \otimes X = \mathrm{Id}$: the *inertia matrix* of X equals the identity matrix. Here $x \otimes x$ is the rank one positive operator on $\mathbf{R}^n$ defined by $(x \otimes x)(y) = \langle x, y\rangle x$. Let (X_i) be i.i.d. copies of X, and consider the random matrices

$$\Gamma_{n,N} = \frac{1}{\sqrt{N}} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix} \tag{4}$$

and, as before

$$A_{n,N} = \Gamma^\dagger_{n,N}\Gamma_{n,N} = \frac{1}{N}\sum_{i=1}^N X_i \otimes X_i.$$

The matrix $A_{n,N}$ is the empirical approximation of the inertia matrix of X, with N sample points. A well-studied class of random vectors is the class of vectors uniformly distributed on a convex body (see the survey [5]). If K is a convex body in $\mathbf{R}^n$ (i.e. a compact full-dimensional convex subset), the random vector X_K is defined as

$$\mathbf{P}(X_K \in B) = \frac{\mathrm{vol}(K \cap B)}{\mathrm{vol}(K)}$$

for any Borel set $B \subset \mathbf{R}^n$. If X_K is isotropic, we say that K is isotropic. It is natural to wonder whether the theorems mentioned before can be extended to large classes of random vectors, especially uniformly distributed on convex bodies. We show that this holds for the simplest examples of convex bodies, the unit balls of the ℓ_p^n spaces, defined as for $1 \leqslant p < +\infty$ by

$$B_p^n = \left\{ (x_1, \dots, x_n) \in \mathbf{R}^n \text{ s.t. } \sum_{i=1}^n |x_i|^p \leqslant 1 \right\}$$

and $B_\infty^n = [-1,1]^n$. We simply write X_p^n instead of $X_{B_p^n}$. The random vector X_p^n can be obtained by simple operations from one-dimensional random variables, as shown by the following theorem due to Ball–Guédon–Mendelson–Naor [3]:

Theorem (Representation of the uniform measure on B_p^n). *Let $1 \leqslant p < +\infty$ and (Y_i) be a n-tuple of i.i.d. random variables distributed according to the probability measure ν_p with density $1/(2\Gamma(1+1/p))e^{-|t|^p}$ $(t \in \mathbf{R})$. Let also Z be an exponential random variable independent from Y (i.e. the density of Z is $e^{-t}, t \geqslant 0$). Then the random vector*

$$\frac{(Y_1, \dots, Y_n)}{(\sum_{i=1}^n |Y_i|^p + Z)^{1/p}}$$

is uniformly distributed on B_p^n.

We write $c_{n,p}$ for the unique positive number such that $c_{n,p}B_p^n$ is isotropic, and write $\tilde{B}_p^n = c_{n,p}B_p^n$:

$$c_{n,p} = \left(\frac{1}{\mathrm{vol}(B_p^n)} \int_{B_p^n} x_1^2 dx_1 \dots dx_n \right)^{-1/2}.$$

We prove the following results

Theorem 1. *Let (n,N) be a sequence of sizes tending to infinity in such a way that the ratio n/N has a limit $\beta \in]0,1[$. Let $A_{n,N} = \Gamma_{n,N}^\dagger \Gamma_{n,N}$ with $\Gamma_{n,N}$ defined as (4), where (X_i) are independent and uniformly distributed on $\tilde{B}_p^n$. Then, almost surely, the sequence of empirical spectral measures $(\mu_{A_{n,N}})$ converges weakly to the Marčenko–Pastur limit $\mu_{(\beta)}$.*

Remark. Theorem 1 has been obtained independently by Pajor and Pastur [7] using the Stieltjes transform method.

Theorem 2. *Let (n,N) be a sequence of sizes tending to infinity in such a way that the ratio n/N has a limit $\beta \in]0,1[$. Let $A_{n,N} = \Gamma_{n,N}^\dagger \Gamma_{n,N}$ with $\Gamma_{n,N}$ defined as (4), where (X_i) are independent and uniformly distributed on $\tilde{B}_p^n$. Then, almost surely,*

$$\lim_{n,N\to\infty} \lambda_{\max}(A_{n,N}) = \lambda_+(\beta),$$

$$\lim_{n,N\to\infty} \lambda_{\min}(A_{n,N}) = \lambda_-(\beta).$$

Remark. The author was able in [1] to adapt the techniques used by Bai–Yin to the setting of random vectors uniformly distributed on unconditional convex bodies, leading to a weaker conclusion (with estimates $1 \pm C\sqrt{\beta}$ instead of $1 \pm \sqrt{\beta}$, for some absolute constant C).

The proof uses the following lemma, which is an immediate consequence of the Barthe–Guédon–Mendelson–Naor representation theorem.

Lemma 1. *Let Γ be defined as (4), where (X_i) are i.i.d. copies of X_p^n. Let $B = (b_{ij})$ be a $N \times n$ random matrix whose entries are independent and distributed according to ν_p. Let Δ be a $N \times N$ diagonal matrix with entries $\delta_{jj} = (\sum_{i=1}^n |b_{ij}|^p + Z_j)^{-1/p}$, where Z_j are i.i.d. exponential random variables independent from B. Then the random matrices Γ and $\frac{c_{n,p}}{\sqrt{N}} \Delta \cdot B$ have the same distribution (here $\cdot$ is the usual matrix product).*

Proof of theorems 1 *and* 2. With notations from lemma 1, the matrices $\Gamma_{n,N}$ and $\frac{c_{n,p}}{\sqrt{N}} \Delta \cdot B$ have the same distribution. We set

$$B' = \frac{1}{\gamma_p} B,$$

$$\Delta' = c_{n,p} \gamma_p \Delta,$$

where $\gamma_p = (\mathbf{E} b_{11}^2)^{1/2}$. We claim that

$$\|\Delta' - \mathrm{Id}\|_{op} \underset{n \to \infty}{\longrightarrow} 0 \quad \text{almost surely.} \tag{5}$$

Using strong versions of the law of large numbers (see [2], Lemma 2) — note that $\mathbf{E}|b_{ij}|^q < \infty$ for all q — we get that, almost surely

$$\lim_{n,N \to \infty} \sup_{1 \leqslant j \leqslant N} \left| \frac{1}{n} \sum_{i=1}^n b_{ij}^p - \mathbf{E} b_{ij}^p \right| = 0.$$

Since $\lim_{n,N \to \infty} \sup_j |\frac{1}{n} Z_j| = 0$ almost surely, we deduce that

$$\lim_{n,N \to \infty} \sup_{1 \leqslant j \leqslant N} \left| (n \mathbf{E}|b_{11}|^p)^{\frac{1}{p}} \delta_{jj} - 1 \right| = 0.$$

Now (5) follows from the fact that

$$c_{n,p} \underset{n \to \infty}{\sim} \frac{n^{1/p} (\mathbf{E}|b_{11}|^p)^{1/p}}{(\mathbf{E} b_{11}^2)^{1/2}}.$$

Consider now $A_{n,N} = \Gamma_{n,N}^\dagger \Gamma_{n,N}$ and $A'_{n,N} = (\frac{1}{\sqrt{N}} B')^\dagger (\frac{1}{\sqrt{N}} B')$. The normalization was chosen so that the matrix B' has independent entries with variance 1, so it enters the setting of the Marčenko–Pastur and the Bai–Yin theorems. We now use the following inequalities, which are proved using the min-max characterization of eigenvalues

$$\frac{1}{\|\Delta'^{-1}\|^2} \lambda_i(A'_{n,N}) \leqslant \lambda_i(A_{n,N}) \leqslant \|\Delta'\|^2 \lambda_i(A'_{n,N}). \tag{6}$$

Theorem 2 is now a direct consequence of (5), (6) and the Bai–Yin theorem applied to the matrices $(A'_{n,N})$. Similarly, the use of Marčenko–Pastur theorem on the matrices $(A'_{n,N})$ gives that for every interval $I \subset \mathbf{R}$, the sequence $(\mu_{A_{n,N}}(I))$ converges almost surely to $\mu_{(\beta)}(I)$. We conclude by arguing that it is enough to test the weak convergence on intervals with rational endpoints (see [4], Theorem 2.2). □

References

[1] G. Aubrun. Sampling convex bodies: a random matrix approach (preprint). `http://www.institut.math.jussieu.fr/~aubrun/sampling.dvi`.

[2] Z. D. Bai and Y. Q. Yin. Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix. *Ann. Probab.*, 21(3):1275–1294, 1993.

[3] F. Barthe, O. Guédon, S. Mendelson, and A. Naor. A probabilistic approach to the geometry of the l_p^n-ball. *Ann. Probab.*, 33(2):480–513, 2005.

[4] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[5] A. Giannopoulos. Notes on isotropic convex bodies (preprint). `http://eudoxos.math.uoa.gr/~apgiannop/isotropic-bodies.ps`, 2003.

[6] V. A. Marčenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72 (114):507–536, 1967.

[7] A. Pajor and L. A. Pastur. Private communication.

Guillaume Aubrun
Institut de Mathématiques de Jussieu, Projet Analyse Fonctionnelle, Université de Paris 6, 175 rue du Chevaleret, 75013 PARIS, France
e-mail: `aubrun@math.jussieu.fr`