Exact simulation of the Marchenko-Pastur distribution

Luc Devroye
School of Computer Science
McGill University

March 13, 2025

Abstract. The Marchenko-Pastur law (Marchenko and Pastur, 1967) describes the limit law of eigenvalues of large rectangular matrices. We give two efficient algorithms for simulating random variables from this distribution.

Keywords and phrases. Random variate generation. Simulation. Monte Carlo method. Expected time analysis. Marchenko-Pastur distribution.

1991 Mathematics Subject Classifications: Primary 65C10. Secondary 65C05, 11K45, 68U20.

Address: School of Computer Science, McGill University, 3480 University Street, Montreal, Canada H3A 2K6. Research sponsored by NSERC Grant A3456. Email: `lucdevroye@gmail.com`

1. Introduction

The Marchenko-Pastur law (Marchenko and Pastur, 1967) describes the limit law of eigenvalues of large rectangular random matrices. In general, it is a mixture of an atomic law with atom at 0 and an absolutely continuous distribution with density given by

$$f(x) = \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi\sigma^2 x}, \lambda_- \le x \le \lambda_+,$$

where $\sigma \in (0, 1]$ is a shape parameter, $\lambda_+ = (1+\sigma)^2$, and $\lambda_- = (1-\sigma)^2$. We will write X_σ for a random variable with this density, and note that $\mathbb{E}\{X_\sigma\} = 1$ and $\mathbb{V}\{X_\sigma\} = \sigma^2$. See, e.g., Bai and Silverstein (2010) for an overview of such limit laws in random matrix theory. The standard description for the Marchenko-Pastur law depends on a parameter $\lambda > 0$: if $\lambda \le 1$, then it is the law of X_σ with $\lambda = \sigma^2$. If $\lambda > 1$, it is a mixed distribution: with probability $1/\lambda$, it is X_1 and with probability $1 - 1/\lambda$, it has an atom at 0.

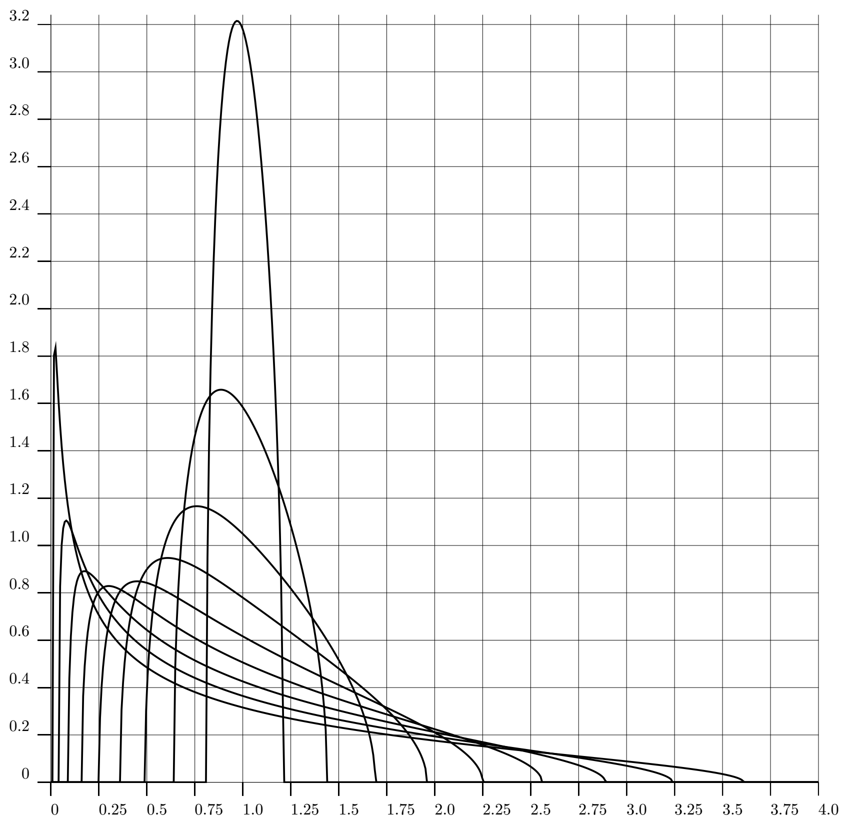

Figure 1. Marchenko-Pastur densities for σ equal to 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 and 0.9, from right to left.

In this note, we point out how one can generate random variates from this distribution.

In their R-package RMTstat, Johnstone, Ma, Perry, Shahram and Biederstedt (2022) used the inverse transformtion method to generate random variates from this distribution. This is necessarily slow and inaccurate if the iterative approximation is halted after a finite number of steps. We offer two alternatives, one suggested by a distributional identity due to Ledoux, and one based on an efficient rejection method.

2. Ledoux's formula

In 2004, Ledoux derived a remarkable decomposition that lends itself quite naturally to simulation. To that end, let $B_{a,b}$ denote a beta random variable with parameters $a, b > 0$. A key role is played by the arcsine law. The arcsine law on $[0, 1]$ is that of a $B_{1/2,1/2}$:

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}, 0 < x < 1.$$

The arcsine random variable ξ on $[-1, 1]$ has density

$$\frac{1}{\pi\sqrt{1-x^2}}, |x| < 1.$$

If $U = B_{1,1}$ denotes a uniform random variable on $[0, 1]$, independent of ξ, then Ledoux obtained the following remarkable representation, adapted for presentation in this note:

$$X_\sigma \stackrel{\mathcal{L}}{=} 1 - \sigma^2 + 2\sigma^2 U + 2\sigma\xi\sqrt{U}\sqrt{1 + \sigma^2(U-1)}. \tag{1}$$

It is immediate from this representation that $\mathbb{E}\{X_\sigma\} = 1$ as ξ is symmetric. Furthermore, as $\sigma \uparrow 1$,

$$X_\sigma \stackrel{\mathcal{L}}{\to} 2U(1+\xi) \stackrel{\mathcal{L}}{=} 4UB_{1/2,1/2} \stackrel{\mathcal{L}}{=} 4B_{1/2,3/2},$$

where $\stackrel{\mathcal{L}}{\to}$ and $\stackrel{\mathcal{L}}{=}$ denote convergence in distribution, and equality in distribution, respectively. This limit behavior is illustrated in Figure 1. As $\sigma \downarrow 0$, $X_\sigma \to 1$ in probability: representation (1) fails to yield a proper nontrivial limit law. However, we observe from (1) that

$$\frac{X_\sigma - 1}{\sigma} \stackrel{\mathcal{L}}{\to} 2\sqrt{U}\xi \stackrel{\mathcal{L}}{=} 4B_{3/2,3/2} - 2,$$

where we note that the latter is Wigner's semicircle law, with density

$$\frac{1}{2\pi}\sqrt{4-x^2}, |x| \le 2.$$

That behavior, too, is illustrated in Figure 1.

Finally, we recall that ξ can be obtained as $\cos(\pi V)$, where V is uniform on $[0, 1]$, or as $N^2/(N^2 + N'^2)$, where N and N' are standard gaussian random variables.

3. A rejection algorithm

It is convenient to write $X_\sigma = \lambda_- + 4\sigma Y$, where Y has density

$$f(y) = \frac{8\sqrt{y(1-y)}}{\pi(\lambda_- + 4\sigma y)}, 0 \le y \le 1.$$

We pick a threshold $\sigma^* \in (0,1)$, and propose using the rejection method (see, e.g., Devroye, 1986) based on these inequalities:

$$f(y) \le \begin{cases} \frac{2}{\pi\sigma}\sqrt{\frac{1-y}{y}} & \text{if } \sigma \ge \sigma^*, \\ \frac{8}{\pi\lambda_-}\sqrt{y(1-y)} & \text{if } \sigma < \sigma^*. \end{cases}$$

We recognize the two limit laws, $B_{1/2,3/2}$, and $B_{3/2,3/2}$ in these bounds. The areas under the envelopes are $1/\sigma$ and $1/\lambda_-$, respectively. The best choice for the threshold is when both areas are equal, which yields one minus the golden ratio,

$$\sigma^* = \frac{3-\sqrt{5}}{2} = \frac{2}{3+\sqrt{5}} = .38196601125010515180\ldots.$$

Applying the rejection algorithm would require an expected number of iterations not exceeding $1/\sigma^* = 2.61803398874989484820\ldots$. More importantly, as σ approaches 0 or 1, the expected number of iterations tends to one, making the method quite efficient.

When $\sigma \ge \sigma^*$, we would generate independent pairs (U,Y) (with U uniform on $[0,1]$ and Y independent and $B_{1/2,3/2}$) until $U(\lambda_- + 4\sigma Y) \le 4\sigma Y$ and return $X_\sigma = \lambda_- + 4\sigma Y$. When $\sigma < \sigma^*$, we would generate independent pairs (U,Y) (with U uniform on $[0,1]$ and Y independent and $B_{3/2,3/2}$) until $U(\lambda_- + 4\sigma Y) \le \lambda_-$ and return $X_\sigma = \lambda_- + 4\sigma Y$. As pointed out earlier, a $B_{1/2,3/2}$ random variable can be obtained as $VB_{1/2,1/2} = V(1+\cos(\pi W))/2$, where V and W are independent uniform $[0,1]$ random variables. Finally, a $B_{3/2,3/2}$ random variable can be generated as $(1+\sqrt{V}\cos(\pi W))/2$.

One by-product of the rejection method developed above is that for both limits as $\sigma \downarrow 0$ and $\sigma \uparrow 1$, the total variation error converges, something that cannot directly be deduced from (1).

All algorithms were tested, with timings of Ledoux's method about 25 percent faster on an iMac machine using the interpreted Postscript language.

4. References

Z. Bai and J. M. Silverstein, *Spectral Analysis of Large Dimensional Random Matrices Second Edition*, Springer, New York, 2010.

L. Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag, New York, 1986.

I. M. Johnstone, Z. Ma, P. O. Perry, M. Shahram, and E. Biederstedt, "The Marcenko-Pastur Distribution," RMTstat: Distributions, Statistics and Tests Derived from Random Matrix Theory, `https://rdrr.io/cran/RMTstat/man/MarcenkoPastur.html`, 2022.

M. Ledoux, “Differential operators and spectral distributions of invariant ensembles from the classical orthogonal polynomials. The continuous case,” *Electronic Journal of Probability*, vol. 9(7), pp. 177–208, 2004.

V. A. Marchenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices (in Russian),” *Sbornik: Mathematics*, vol. 1, pp. 457–483, 1967.