What is the expected number of points drawn from a Determinantal Point Process defined by a Wishart kernel?

Josefina Correa

Abstract

Determinantal Point Processes(DPPs) define probability measures over the power set of $1 : \mathcal{N}$ outcomes. Random samples from a DPP can be obtained by using the eigendecomposition of a positive semi-definite matrix $\mathcal{L}$. However, the how this positive semi-definite matrix $\mathcal{L}$ affects the expected number of points remains unknown. This project will study the distribution of the number of outcomes from a DPP defined by random $\mathcal{L}$ matrices, with a focus on Wishart matrices.

1 Introduction

Given a discrete space $\mathcal{X}$, a stochastic point process $\mathcal{P}$ on $\mathcal{X}$ is defined as a probability measure on the power set of $\mathcal{X}$. If there exists a matrix $\mathcal{L}$ such that the probability of observing all possible subsets $\mathcal{I}$, of $\mathcal{X}$, is given by the determinant of the principal minor of $\mathcal{L}$, then $\mathcal{P}$ is a determinantal point process. Since every principal minor is non-negative, $\mathcal{L}$ is positive semi-definite. An alternative way of assigning probabilities on subsets $\mathcal{I}$ is by using the ”L” method, also denoted ”L-ensembles” [1], by using a positive semi-definite matrix $\mathcal{L}$

$$\mathbb{P}(\mathcal{I}) = \frac{\mathcal{L}\binom{\mathcal{I}}{\mathcal{I}}}{\det\left(\mathcal{I} + \mathcal{L}\right)} \tag{1}$$

Here, $\mathcal{I}$ is a subset of $\mathcal{X}$ This project studied $p_X(x)$ and $p_{\tilde{X}}(x)$, the distribution of the expected number of points drawn from a DPP defined by real and compelx Wishart matrices, denoted as $W(m, n)$ and $\tilde{W}(m, n)$, respectively. Theoretical and numerical approximations for these distributions were derived using (1) random matrix theory and (2) monte carlo simulations. [3, 2, 1]

2 Methods

By noticing that the expected number of points, X is the sum of non-identically distributed bernoulli random variables, $\mathcal{I}_k$, we can arrive at the probability of observing k counts, for $k \in (0, K)$. This distribution is also known as the Poisson-Binomial distribution. [4]

$$\begin{aligned}
\mathcal{I}_k &\sim \text{Bernoulli}(\lambda_k/(1+\lambda_k)), \lambda_k \text{ is an eigenvalue of } \mathcal{L} \\
X &= \sum_{m=1}^{M} \mathcal{I}_m, \\
\mathcal{P} &=: \text{ powerset of 1:M}, \\
\mathcal{P}_k &=: \text{ set of all p } \in \mathcal{P} \text{ with length k} \\
\mathbb{P}(X = k) &= \sum_{J \in \mathcal{P}_k} \prod_{i \in J} \frac{\lambda_i}{1+\lambda_i} \cdot \prod_{j \in J^c} \left(1 - \frac{\lambda_j}{1+\lambda_j}\right) \\
&= \sum_{J \in \mathcal{P}_k} \prod_{i \in J} \frac{\lambda_i}{1+\lambda_i} \cdot \prod_{j \in J^c} \left(\frac{1}{1+\lambda_j}\right) \\
&= \frac{1}{\prod_{m=1}^{M}(1+\lambda_m)} \cdot \sum_{J \in \mathcal{P}_k} \prod_{i \in J} \lambda_i
\end{aligned} \tag{2}$$

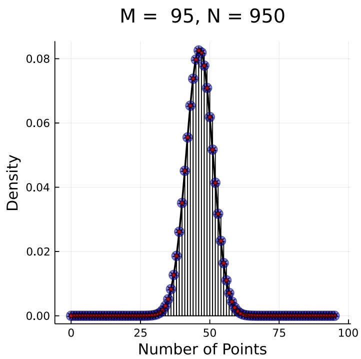
Re-writing this PMF in terms of the coefficients of Laguerre polynomials, with $\alpha = n - m$, the number of columns of a random gaussian marix minus its number of rows, we get the following expression:

$$\begin{aligned}
W(m,n) &= AA'/n \ , \ A_{i,j} \overset{\text{iid}}{\sim} N(0,1) \ , \ k \in 0 : m \\
\mathcal{L}_m^{(\alpha)}(x) &= \sum_{i=0}^{m} \frac{(-1)^i}{i!} \binom{m+\alpha}{m-i} x^i \cdot \frac{1}{n^{m-i}} \\
\mathbb{P}(X=k) &= \frac{1}{\prod_{m=1}^{M}(1+\lambda_m)} \cdot \sum_{J \in \mathcal{P}_k} \prod_{i \in J} \lambda_i \\
\mathbb{P}(X=k) &= \frac{1}{\mathcal{L}_m^{(n-m)}(-1)} \cdot \frac{1}{(m-k)!} \binom{m+\alpha}{k} \cdot \frac{1}{n^k} \qquad (3)
\end{aligned}$$

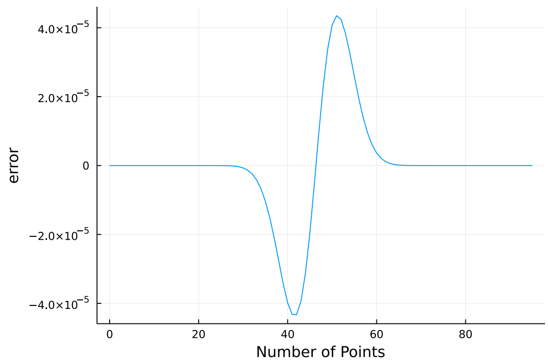
Equation (3) contains a factor $\frac{1}{n^k}$ that accounts for the fact that in the Wishart case, the eigenvalues are scaled by the number of rows in the random gaussian matrix of size $m \times n$. The code for testing both approaches is included in a Pluto notebook. Gaussian approximations were computed by taking $\mu = \sum_{m=1}^{M} \frac{\lambda_m}{1+\lambda_m}$ and $\sigma^2 = \sum_{m=1}^{M} \frac{\lambda_m}{1+\lambda_m}(1 - \frac{\lambda_m}{1+\lambda_m})$.

3 Results

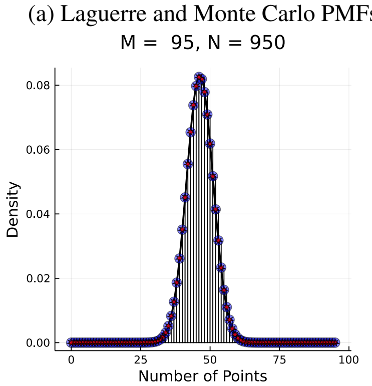
For different values of $r = M/N$, we can see that the Laguerre and Monte Carlo approximations are in agreement. Figures 1 through 4 show results for numerical experiments, using different values of r. The errors, computed as the difference between the Laguerre and Monte Carlo approximations, range from -10^-5 to 10^-5. Interestingly, for a rank 1 Wishart matrix, both approaches yield reasonable answers. The density in this case has a value close to 1 at 1 point, a value close to 0 at 0, and 0, elsewhere. This is reasonable, since a rank 1 matrix will have one eigenvalue equal to its trace, and all other eigenvalues equal to 0. Therefore, when constructing probabilities as $\lambda_i/(1+\lambda_i)$, only one of these should be above 0.5, hence the expected number of outcomes should be 1. It is worth noting that there are various ways of setting up the Monte Carlo simulation. One may construct an average characteristic polynomial, and estimate the PMF by normalizing its coefficients by the sum of the absolute value of the coefficients. One may also, at each iteration, construct a PMF, and then average over all Monte Carlo PMFs. Both approaches yield similar answers, though the former produces an estimate closer to that obtained analytically. Results are shown using the first approach (using the average characteristic polynomial).



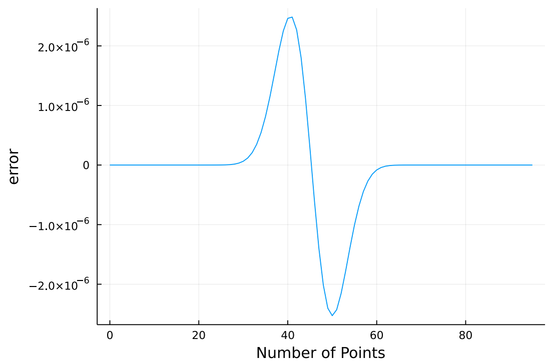
(a) Laguerre and Monte Carlo PMFs for $\beta = 1$



(b) Monte Carlo approximation error for $\beta = 1$

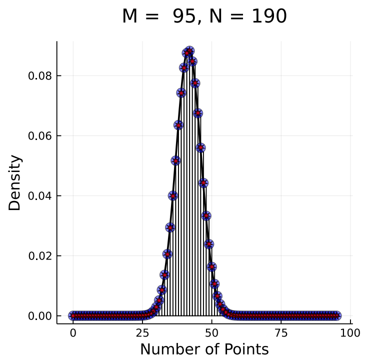


(c) Laguerre and Monte Carlo PMFs for $\beta = 2$

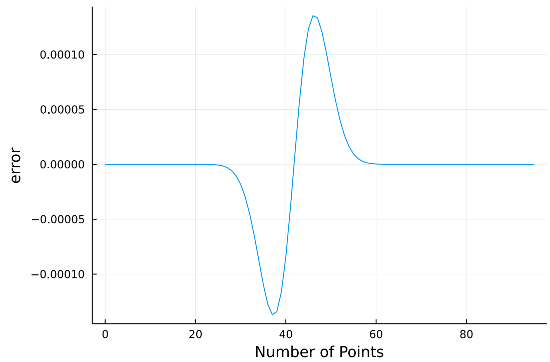


(d) Monte Carlo approximation error for $\beta = 2$

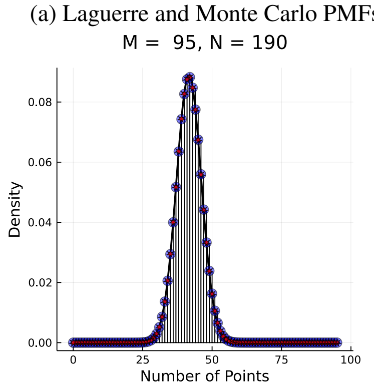
Figure 1: Laguerre (blue circles) and Monte Carlo (red stars) approximations for r=0.1. Black line shows a gaussian approximation to this density.



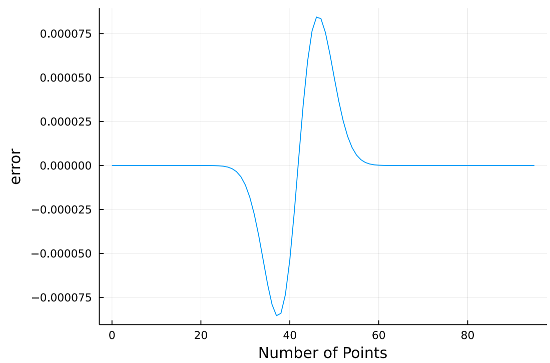
(a) Laguerre and Monte Carlo PMFs for $\beta = 1$



(b) Monte Carlo approximation error for $\beta = 1$

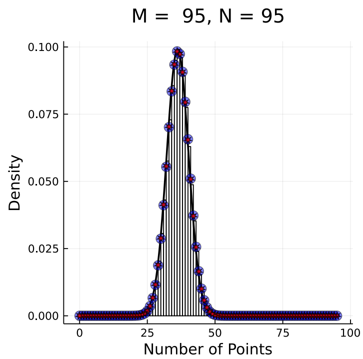


(c) Laguerre and Monte Carlo PMFs for $\beta = 2$

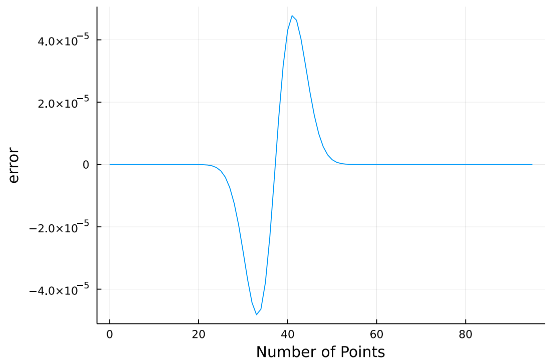


(d) Monte Carlo approximation error for $\beta = 2$

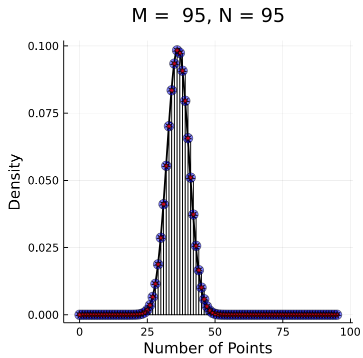
Figure 2: Laguerre (blue circles) and Monte Carlo (red stars) approximations for r=0.5. Black line shows a gaussian approximation to this density.



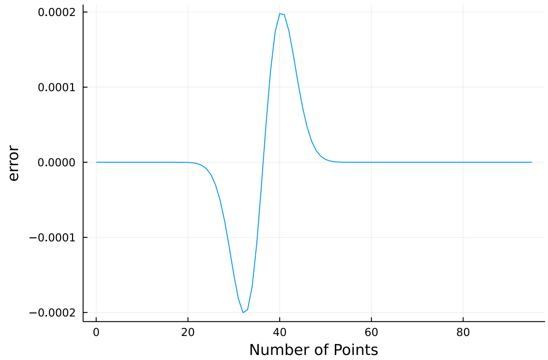
(a) Laguerre and Monte Carlo PMFs for $\beta = 1$



(b) Monte Carlo approximation error for $\beta = 1$

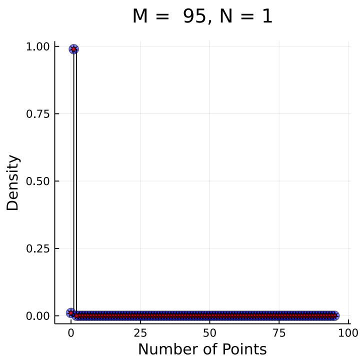


(c) Laguerre and Monte Carlo PMFs for $\beta = 2$

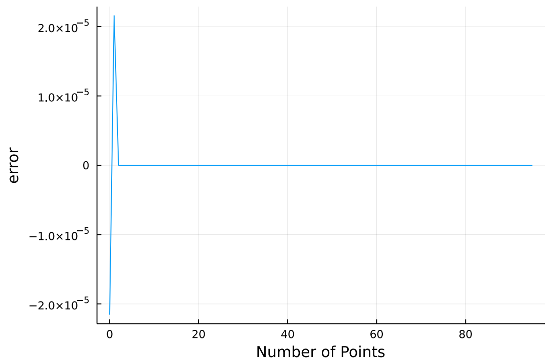


(d) Monte Carlo approximation error for $\beta = 2$

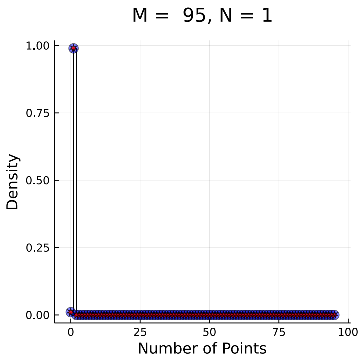
Figure 3: Laguerre (blue circles) and Monte Carlo (red stars) approximations for r=1.0. Black line shows a gaussian approximation to this density.



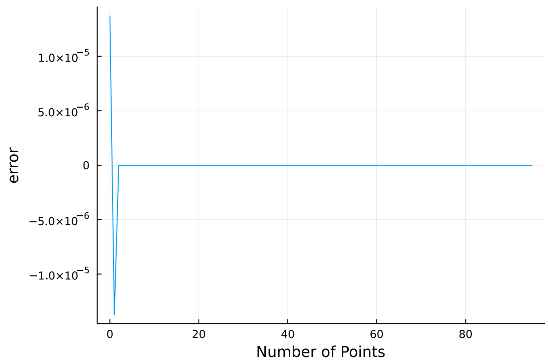
(a) Laguerre and Monte Carlo PMFs for $\beta = 1$



(b) Monte Carlo approximation error for $\beta = 1$



(c) Laguerre and Monte Carlo PMFs for $\beta = 2$



(d) Monte Carlo approximation error for $\beta = 2$

Figure 4: Laguerre (blue circles) and Monte Carlo (red stars) approximations for r=m (rank 1 case). Black line shows a gaussian approximation to this density.

4 Concluding remarks and future directions

A determinantal point process may be defined in terms of a positive semi-definite matrix. This project explored the Wishart case, both for hermitian and symmetric Wishart matrices. It would be interesting to explore whether connections can be made to other kinds of random matrices, such as Jacobi or Hermite matrices. The Hermite case might be more challenging as it would require generating a random matrix (A+A)'/2 such that all eigenvalues are positive. For the Jacobi matrices, this would require looking into closed-form expressions for the coefficients of Jacobi polynomials.

References

[1] Alexei Borodin. "Determinantal point processes". In: *arXiv preprint arXiv:0911.1153* (2009).

[2] Alan Edelman. "Determinantal Point Process Notes (18.338)". In: *Unpublished Work* (2021).

[3] Alan Edelman. "Eigenvalues and condition numbers of random matrices". PhD thesis. Cambridge, MA, 1989.

[4] Joseph L Hodges and Lucien Le Cam. "The Poisson approximation to the Poisson binomial distribution". In: *The Annals of Mathematical Statistics* 31.3 (1960), pp. 737–740.