Determinantal point process: the spherical ensemble

Joaquim Ortega-Cerdà

Barcelona, March 4 2015

UNIVERSITAT DE BARCELONA

Determinantal point process

Definition

A determinantal point process A is a random point process such that the joint intensities have the form:

$$\rho_n(x_1, \ldots, x_n) = det(K(x_i, x_j)_{i,j \leq n}).$$

Determinantal point process

Definition

A determinantal point process A is a random point process such that the joint intensities have the form:

$$\rho_n(x_1, \ldots, x_n) = det(K(x_i, x_j)_{i,j \leq n}).$$

Recall that the joint intensities ρ_k satisfy:

$$\mathbb{E} \sum_{x_1,\ldots,x_k \in A} f(x_1, \ldots, x_k) = \int f(x_i, \ldots, x_k) \rho_k(x_i, \ldots, x_k)$$

for any f symmetric bounded and of compact support.

Determinantal process: the origin

They were systematically studied at the end of 70's by Macchi because they are a good model for fermions:

Determinantal process: the origin

They were systematically studied at the end of 70's by Macchi because they are a good model for fermions:
In quantum mechanics the position of a particle is represented by a wave function ψ such that $\int |\psi|^2 = 1$. If we have n independent particles, the global system is defined by a global wave functionn $\Psi(x_1, \ldots, x_n)$.

Determinantal process: the origin

They were systematically studied at the end of 70's by Macchi because they are a good model for fermions:
In quantum mechanics the position of a particle is represented by a wave function ψ such that $\int |\psi|^2 = 1$. If we have n independent particles, the global system is defined by a global wave functionn $\Psi(x_1, \ldots, x_n)$. If we want to model fermions by Pauli exclusion principle the composite wave function must be antysimetric in the different variables. This suggests

$$\Psi = c \det(\psi_i(x_j))$$

Then

$$|\Psi|^2 = c_N^2 \det(K(x_i, x_j))$$

where $K(x,y) = \sum \psi_i(x)\overline{\psi}_i(y)$

General facts

If the point process has n points almost surely then the kernel K defines an integral operator: the orthogonal projection onto a subspace of L^2 of dimension n.

General facts

If the point process has n points almost surely then the kernel K defines an integral operator: the orthogonal projection onto a subspace of L^2 of dimension n.
In general

Theorem (Macchi, Soshnikov)

An hermitic kernel $K(x, y)$ corresponds to a determinantal point process if and only if the integral operator $T : L^2 \to L^2$ has all eigenvalues $\lambda \in [0, 1]$.

General facts

If the point process has n points almost surely then the kernel K defines an integral operator: the orthogonal projection onto a subspace of L^2 of dimension n.

In general

Theorem (Macchi, Soshnikov)

An hermitic kernel $K(x, y)$ corresponds to a determinantal point process if and only if the integral operator $T : L^2 \to L^2$ has all eigenvalues $\lambda \in [0, 1]$.

Moreover:

Theorem (Shirai, Takahashi)

In a determinantal process, the number of points that fall in a compact set D has the same distribution as a sum of independent Bernoulli(λ_i^D)) random variables where λ_i^D are the eigenvalues of the operator T restricted to D.

Determinantal processes everywhere

This algebraic structure is very prevalent. Some notorious examples of determinantal point processes:

- Some physical models of fermions. (Macchi, Soshnikov)

Determinantal processes everywhere

This algebraic structure is very prevalent. Some notorious examples of determinantal point processes:

- Some physical models of fermions. (Macchi, Soshnikov)
- The eigenvalues of random matrices: The GUE ensemble, the CUE ensemble, the Ginibre ensemble (Wigner, Weyl and Dyson, Ginibre)

Determinantal processes everywhere

This algebraic structure is very prevalent. Some notorious examples of determinantal point processes:

- Some physical models of fermions. (Macchi, Soshnikov)
- The eigenvalues of random matrices: The GUE ensemble, the CUE ensemble, the Ginibre ensemble (Wigner, Weyl and Dyson, Ginibre)
- Zeros of random polynomials of type $\sum a_n z^n$ where a_n are i.i.d. random variables with distribution $N_{\mathbb{C}}(0, 1)$. (Peres and Virag)

Determinantal processes everywhere

This algebraic structure is very prevalent. Some notorious examples of determinantal point processes:

- Some physical models of fermions. (Macchi, Soshnikov)
- The eigenvalues of random matrices: The GUE ensemble, the CUE ensemble, the Ginibre ensemble (Wigner, Weyl and Dyson, Ginibre)
- Zeros of random polynomials of type $\sum a_n z^n$ where a_n are i.i.d. random variables with distribution $N_{\mathbb{C}}(0,1)$. (Peres and Virag)
- Non-intersecting random walks (Karlin and Mc Gregor)

Determinantal processes everywhere

This algebraic structure is very prevalent. Some notorious examples of determinantal point processes:

- Some physical models of fermions. (Macchi, Soshnikov)
- The eigenvalues of random matrices: The GUE ensemble, the CUE ensemble, the Ginibre ensemble (Wigner, Weyl and Dyson, Ginibre)
- Zeros of random polynomials of type $\sum a_n z^n$ where a_n are i.i.d. random variables with distribution $N_{\mathbb{C}}(0,1)$. (Peres and Virag)
- Non-intersecting random walks (Karlin and Mc Gregor)
- Uniform spanning trees (Burton and Pemantle) and (Benjamini, Lyons, Peres and Schramm)

Random matrices, the Ginibre ensemble

The Ginibre ensemble is the spectrum of an $n \times n$ random matrix with i.i.d entries with $N_{\mathbb{C}}(0, 1)$ entries. The joint distribution of the eigenvalues is given by the law:

$$\prod_{k=1}^{N} \frac{1}{k!} \prod_{i=1}^{N} e^{-|z_i|^2} \prod_{j<k} |z_j - z_k|^2 dm(z)$$

Random matrices, the Ginibre ensemble

The Ginibre ensemble is the spectrum of an $n \times n$ random matrix with i.i.d entries with $N_{\mathbb{C}}(0,1)$ entries. The joint distribution of the eigenvalues is given by the law:

$$\prod_{k=1}^{N} \frac{1}{k!} \prod_{i=1}^{N} e^{-|z_i|^2} \prod_{j<k} |z_j - z_k|^2 dm(z)$$

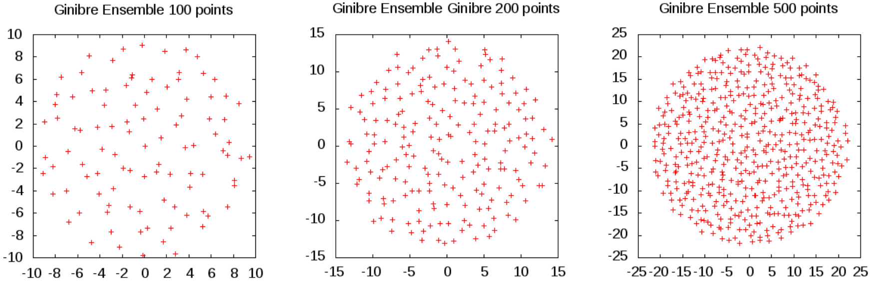
This is determinantal if we observe

$$\prod_{i=1}^{N} \frac{e^{-|z_i|^2}}{i!} \prod_{j<k} |z_j - z_k|^2 = |det(z_i^j \frac{e^{-|z_i|^2/2}}{\sqrt{i!}})|^2 = \det(K(z_i, z_j))$$

where $K(z,w) = \sum_k \frac{(z\overline{w})^k}{k!} e^{-|z|^2/2} e^{-|w|^2/2}$.

Ginibre ensemble

A typical instance is:



Thus, typically the eigenvalues lie in a disk of radious $\sqrt{n}$. If one rescales $\sqrt{n}$ Ginibre proved that almost surely:

$$\frac{1}{n}\sum_i \delta_{(\lambda_i/\sqrt{n})} \overset{*}{\rightharpoonup} \frac{1}{\pi}\chi_{D(0,1)}$$

Spherical ensembles

Krishnapur considered the following point process: Let A, B be n by n random matrices with i.i.d. Gaussian entries. Then he proved that the generalized eigenvalues associated to the pair (A, B), i.e. the eigenvalues of $A^{-1}B$ have joint probability density (wrt Lebesgue measure):

$$C_n \prod_{k=1}^{n} \frac{1}{(1+|z_k|^2)^{n+1}} \prod_{i<j} |z_i - z_j|^2.$$

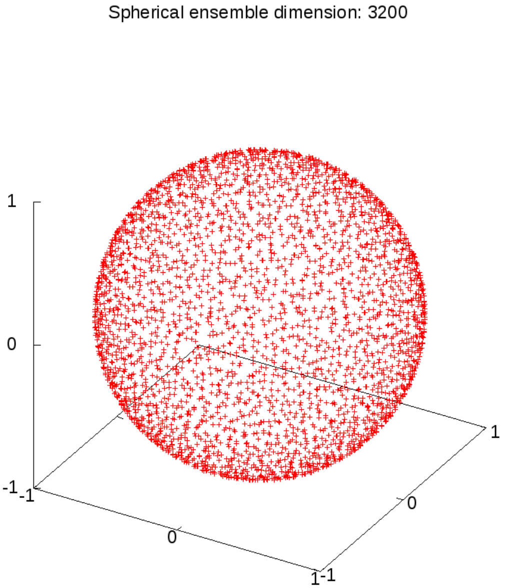
Spherical ensembles

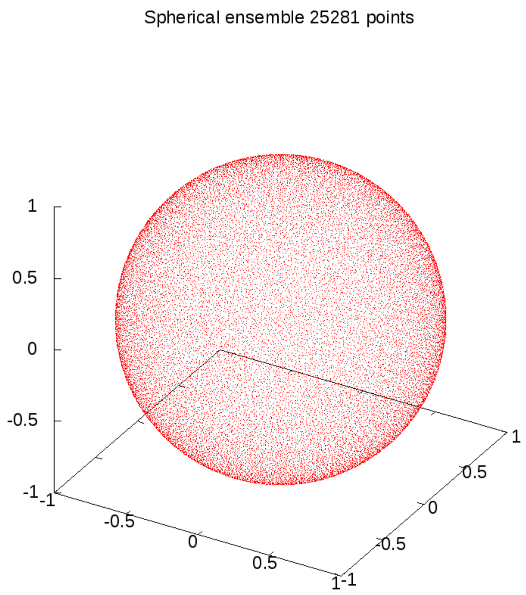
Krishnapur considered the following point process: Let A, B be n by n random matrices with i.i.d. Gaussian entries. Then he proved that the generalized eigenvalues associated to the pair (A, B), i.e. the eigenvalues of $A^{-1}B$ have joint probability density (wrt Lebesgue measure):

$$C_n \prod_{k=1}^{n} \frac{1}{(1+|z_k|^2)^{n+1}} \prod_{i<j} |z_i - z_j|^2.$$

If we consider the stereographic projection to the sphere $\mathbb{S}^2$, then the joint density (with respect to the product area measure in the sphere) is

$$K_n \prod_{i<j} \|P_i - P_j\|^2_{\mathbb{R}^3}.$$


Spherical ensemble dimension: 3200
1
0
-1
-1
0
1
-1
0
1


Spherical ensemble 25281 points
1
0.5
0
-0.5
-1
-1
-0.5
0
0.5
1
-1
-0.5
0
0.5
1

The space of functions

Let P_n be the space functions defined as

$$q(z) = \frac{p(z)}{(1+|z|^2)^{(n-1)/2}},$$

where p is a polynomial of degree less than n. Clearly $P_n \subset L^2(\mu)$, where $d\mu(z) = 1/(1+|z|^2)^2$. It is a reproducing kernel Hilbert space. Its reproducing kernel is

$$K_n(z,w) = \frac{(1+z\bar{w})^{n-1}}{(1+|z|^2)^{(n-1)/2}(1+|w|^2)^{(n-1)/2}}$$

.

A determinantal form

We have that the matrix

$$\begin{pmatrix} \overline{q_1(z_1)} & \cdots & \overline{q_n(z_1)} \\ \vdots & \ddots & \vdots \\ \overline{q_1(z_n)} & \cdots & \overline{q_n(z_n)} \end{pmatrix} \begin{pmatrix} q_1(z_1) & \cdots & q_1(z_n) \\ \vdots & \ddots & \vdots \\ q_n(z_1) & \cdots & q_n(z_n) \end{pmatrix} = \begin{pmatrix} K_n(z_1,z_1) & \cdots & K_n(z_1,z_n) \\ \vdots & \ddots & \vdots \\ K_n(z_n,z_1) & \cdots & K_n(z_n,z_n) \end{pmatrix}$$

A determinantal form

We have that the matrix

$$\begin{pmatrix} \overline{q_1(z_1)} & \cdots & \overline{q_n(z_1)} \\ \vdots & \ddots & \vdots \\ \overline{q_1(z_n)} & \cdots & \overline{q_n(z_n)} \end{pmatrix} \begin{pmatrix} q_1(z_1) & \cdots & q_1(z_n) \\ \vdots & \ddots & \vdots \\ q_n(z_1) & \cdots & q_n(z_n) \end{pmatrix} = \begin{pmatrix} K_n(z_1,z_1) & \cdots & K_n(z_1,z_n) \\ \vdots & \ddots & \vdots \\ K_n(z_n,z_1) & \cdots & K_n(z_n,z_n) \end{pmatrix}$$

Thus

$$\begin{vmatrix} K_n(z_1, z_1) & \cdots & K_n(z_1, z_n) \\ \vdots & & \vdots \\ K_n(z_n, z_1) & \cdots & K_n(z_n, z_n) \end{vmatrix} = \begin{vmatrix} q_1(z_1) & \cdots & q_1(z_n) \\ \vdots & & \vdots \\ q_n(z_1) & \cdots & q_n(z_n) \end{vmatrix}^2$$

Therefore the spherical ensemble generates a *determinantal* point process.

Weak convergence of empirical measure

Given a realization $z_1, \ldots, z_n$ of the random point process we denote by $\mu_n = \frac{1}{n}\sum_i \delta_{z_i}$ to the empirical measure. We take a sequence μ_n, $n = 1, 2, \ldots$ of independent point process of the spherical ensemble. The normalized measure on the sphere is denoted by μ.

Weak convergence of empirical measure

Given a realization $z_1, \ldots, z_n$ of the random point process we denote by $\mu_n = \frac{1}{n}\sum_i \delta_{z_i}$ to the empirical measure. We take a sequence μ_n, $n = 1, 2, \ldots$ of independent point process of the spherical ensemble. The normalized measure on the sphere is denoted by μ.

Theorem

With probability one $\mu_n \overset{*}{\rightharpoonup} \mu$. More precisely the Kantorovich-Wasserstein distance $KW_1(\mu_n, \mu) \lesssim \frac{\log n}{\sqrt{n}}$ with probability one.

The Kantorovich-Wasserstein distance

Given a compact metric space K we defines the KW_1 distance between two probability measures μ and ν supported in K as

$$KW_1(\mu, \nu) = \inf_{\rho} \iint_{K \times K} d(x, y) d\rho(x, y),$$

where ρ is an admissible probability measure, i.e. the marginals of ρ are μ and ν respectively.

The Kantorovich-Wasserstein distance

Given a compact metric space K we defines the KW_1 distance between two probability measures μ and ν supported in K as

$$KW_1(\mu, \nu) = \inf_{\rho} \iint_{K \times K} d(x, y) d\rho(x, y),$$

where ρ is an admissible probability measure, i.e. the marginals of ρ are μ and ν respectively. Alternatively:

$$KW_1(\mu, \nu) = \inf_{\rho} \iint_{K \times K} d(x, y) d|\rho|(x, y),$$

where ρ is an admissible complex measure, i.e. the marginals of ρ are μ and ν respectively

The Lagrange functions

Given any sequence of points $(z_1, \ldots, z_n)$ we define the Lagrange functions:

$$\ell_j(z) = \frac{\begin{vmatrix} q_1(z_1) & \cdots & q_1(z) & \cdots & q_1(z_n) \\ \vdots & & \vdots & & \vdots \\ q_n(z_1) & \cdots & q_n(z) & \cdots & q_n(z_n) \end{vmatrix}}{\begin{vmatrix} q_1(z_1) & \cdots & q_1(z_j) & \cdots & q_1(z_n) \\ \vdots & & \vdots & & \vdots \\ q_n(z_1) & \cdots & q_n(z_j) & \cdots & q_n(z_n) \end{vmatrix}}$$

Clearly $\ell_j \in P_n$ and $\ell_j(z_i) = 0$ if $i \neq j$ and $\ell_j(z_j) = 1$.

Lagrange functions and the density function

We have that the joint distribution of $(z_1, \ldots, z_n)$ is given by the density (with respect to the spherical measure $d\mu(z_1) \cdots d\mu(z_n)$):

$$\rho_n(z_1, \ldots, z_n) = \begin{vmatrix} K_n(z_1, z_1) & \cdots & K_n(z_1, z_n) \\ \vdots & & \vdots \\ K_n(z_n, z_1) & \cdots & K_n(z_n, z_n) \end{vmatrix},$$

Lagrange functions and the density function

We have that the joint distribution of $(z_1, \dots, z_n)$ is given by the density (with respect to the spherical measure $d\mu(z_1) \cdots d\mu(z_n)$):

$$\rho_n(z_1, \dots, z_n) = \begin{vmatrix} K_n(z_1, z_1) & \cdots & K_n(z_1, z_n) \\ \vdots & & \vdots \\ K_n(z_n, z_1) & \cdots & K_n(z_n, z_n) \end{vmatrix},$$

and

$$|\ell_j(z)|^2 = \frac{\rho_n(z_1, \dots, z, \dots, z_n)}{\rho_n(z_1, \dots, z_j, \dots, z_n)}.$$

The transport plan

Consider the transport plan

$$p(z,w) = \frac{1}{n}\sum_{j=1}^{n} \delta_{z_j}(w) K_n(z,z_j) \ell_j(z)\, d\mu(z).$$

It has the right marginals $\frac{1}{n}\sum \delta_{z_j}$ and μ respectively and thus

$$KW_1(\mu_n,\mu) \leq \iint |z-w| d|p| \leq \frac{1}{n}\sum_{j=1}^{n} \int d(z,z_j)|\ell_j(z)||K_n(z,z_j)| d\mu(z).$$

Estimating the K-W distance

$$(\mathbb{E}KW_1)^2 \leq$$

$$\int_{\mathbb{C}^n} \frac{1}{n} \sum_{j=1}^{n} \left(\int_{\mathbb{C}} d(z, z_j) |\ell_j(z)| |K_n(z, z_j)| d\mu(z) \right)^2 \frac{\rho_n(z_1, \ldots, z_n)}{n!} \leq$$

$$\int_{\mathbb{C}^n} \frac{1}{n} \sum_{j=1}^{n} \left(\int_{\mathbb{C}} d(z, z_j)^2 |K_n(z, z_j)| \int_{\mathbb{C}} |\ell_j(z)|^2 |K_n(z, z_j)| \right) \frac{\rho_n(z_1, \ldots, z_n)}{n!}.$$

Off diagonal decay of the reproducing kernel

It is easy to see that

$$\sup_{w\in\mathbb{C}} \int_{\mathbb{C}} |z-w|^2 |K_n(z,w)| d\mu(z) \le C/n,$$

and

$$\sup_{w\in\mathbb{C}} \int_{\mathbb{C}} |K_n(z,w)| d\mu(z) \le C.$$

Off diagonal decay of the reproducing kernel

It is easy to see that

$$\sup_{w\in\mathbb{C}} \int_{\mathbb{C}} |z-w|^2 |K_n(z,w)| d\mu(z) \leq C/n,$$

and

$$\sup_{w\in\mathbb{C}} \int_{\mathbb{C}} |K_n(z,w)| d\mu(z) \leq C.$$

because

$$\begin{aligned} |K_n(z,w)|^2 =& n^2 \left(1 - \frac{|z-w|^2}{(1+|z|^2)(1+|w|^2)}\right)^{n-1} \leq \\ \leq& K n^2 \exp\left(-Cn \frac{|z-w|^2}{(1+|z|^2)(1+|w|^2)}\right). \end{aligned}$$

The final estimate

$$(\mathbb{E}KW_1)^2 \le \frac{C}{n}\int_{\mathbb{C}^n}\frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{C}}|\ell_j(z)|^2|K_n(z,z_j)|\frac{\rho_n(z_1,\dots,z_j,\dots,z_n)}{n!} =$$

$$\frac{C}{n}\int_{\mathbb{C}^n}\frac{1}{n}\sum_{j=1}^{n}\int_{\mathbb{C}}|K_n(z,z_j)|\frac{\rho_n(z_1,\dots,z,\dots,z_n)}{n!} \le \frac{C}{n}.$$

A concentration of measure

We want to study now the empirical measure. For determinantal process we have:

Theorem (Pemantle-Peres)

Let Z be a determinantal point process of n points. Let f be a Lipschitz-1 functional on finite counting measures (with respect to the total variation distance). Then

$$\mathbb{P}(f - \mathbb{E}f \geq a) \leq 3 \exp\left(-\frac{a^2}{16(a+2n)}\right)$$

A concentration of measure

We want to study now the empirical measure. For determinantal process we have:

Theorem (Pemantle-Peres)

Let Z be a determinantal point process of n points. Let f be a Lipschitz-1 functional on finite counting measures (with respect to the total variation distance). Then

$$\mathbb{P}(f - \mathbb{E}f \geq a) \leq 3 \exp\left(-\frac{a^2}{16(a+2n)}\right)$$

The functional $f(\sigma) = nKW_1(\frac{1}{n}\sigma, \mu)$ is Lipshchitz-1.

Almost sure convergence

To finish take $a = 10\sqrt{n\log(n)}$, then

$$\mathbb{P}\Big(KW_1(\mu_n,\mu) > \frac{11\sqrt{\log(n)}}{\sqrt{n}}\Big) \leq$$

$$3\exp\left(-\frac{100n\log(n)}{16(10\sqrt{n\log(n)}+2n)}\right) \lesssim \frac{1}{n^2}.$$

Almost sure convergence

To finish take $a = 10\sqrt{n\log(n)}$, then

$$\mathbb{P}\Big(KW_1(\mu_n,\mu) > \frac{11\sqrt{\log(n)}}{\sqrt{n}}\Big) \leq$$

$$3\exp\left(-\frac{100n\log(n)}{16(10\sqrt{n\log(n)}+2n)}\right) \lesssim \frac{1}{n^2}.$$

Now a standard application of the Borel-Cantelli lemma shows that with probability one

$$KW_1(\mu_n,\mu) \leq \frac{10\sqrt{\log n}}{\sqrt{n}}.$$

Application

There is an open question by Smale: Can one find $X = (x_1, \dots, x_N) \subset \mathbb{S}^2$ such that $\mathcal{E}(X) - m_N \leq c \log N$, c a universal constant?
Here $\mathcal{E}(X) = -\sum_{i \neq j} \log \|x_i - x_j\|$ and $m_N = \min_{X \subset \mathbb{S}^2} \mathcal{E}(X)$.

Application

There is an open question by Smale: Can one find $X = (x_1, \ldots, x_N) \subset \mathbb{S}^2$ such that $\mathcal{E}(X) - m_N \leq c \log N$, c a universal constant?
Here $\mathcal{E}(X) = -\sum_{i \neq j} \log \|x_i - x_j\|$ and $m_N = \min_{X \subset \mathbb{S}^2} \mathcal{E}(X)$.
The best known estimates of m_N are

$$m_N = (\frac{1}{2} - \log 2)N^2 - \frac{N \log N}{2} + C_N N,$$

where $-0.22553754 \leq \liminf C_N \leq \limsup C_N \leq -0.0469945$.
The spherical ensemble provides a good candidate as they can be constructed with N^3 operations.

Theorem (Alishahi and Zamani)

$$\mathbb{E}\mathcal{E}(P_1, \ldots, P_N) =$$
$$= \left(\frac{1}{2} - \log 2\right) N^2 - \frac{1}{2} N \log N + \left(\log 2 - \frac{\gamma}{2}\right) N - \frac{1}{4} + O\left(\frac{1}{N}\right)$$

Here, γ is the Euler constant.

This is close to the best known estimates.