Matrices de Fourier aléatoires et estimations spectrales non asymptotiques

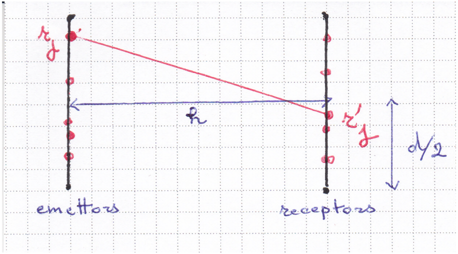
Aline Bonami,
Université d'Orléans, France

travail en collaboration avec Abderrazek Karoui
Université de Carthage, Tunisie

Journée de l'institut Denis Poisson,
le 14 juin 2018

Transmission in a wireless MIMO network.

Considered by Desgroseilliers, Lévêque and Preissmann, EPFL (2013) :



Matrix used to describe the *phase fading* between the emission and the reception :

$$a_{jk} = \frac{e^{2i\pi|r_j - r'_k|/\lambda}}{|r_j - r'_k|}.$$

Here r_j is the position of the $j-$th emitter node while r'_k is the poisition of the $k-$th receptor node. Both $n-$tuples of nodes are chosen randomly in two intervals and at large distance.

Their questions on the channel fading matrix.

$$a_{jk} = \frac{e^{2i\pi|r_j - r'_k|/\lambda}}{|r_j - r'_k|}.$$

- Number of degrees of freedom, that is, the number of significative singular values (square roots of eigenvalues of A^*A).

Their questions on the channel fading matrix.

$$a_{jk} = \frac{e^{2i\pi|r_j - r'_k|/\lambda}}{|r_j - r'_k|}.$$

- Number of degrees of freedom, that is, the number of significative singular values (square roots of eigenvalues of A^*A).
- For a value p, which is the total power of the input signal, uniformly distributed among the n nodes, want to have an approximation of the Shannon capacity of the system, given by

$$C(p) = \log_2(\det(I + \frac{p}{n}A^*A)) := \sum_j \log_2(1 + \frac{p}{n}\lambda_j(A)^2).$$

Nodes are chosen randomly. We want to know this with large probability.

The model (Desgroseilliers et al.).

$$a_{jk} = \frac{e^{2i\pi|r_j - r'_k|/\lambda}}{|r_j - r'_k|}.$$

They propose to replace $|r_j - r'_k|$ in the numerator by the first terms in the development in $1/h$ and the denominator by h.

$$|r_j - r'_k| \approx h + \frac{d^2(y'_k - y_j)^2}{2h}.$$

Here y_j and y'_k belong to $(-1/2, +1/2)$ and dy_j, dy'_k are the second coordinates of the points r_j, r'_k.
We pose $m := \frac{d^2}{h\lambda}$ (λ is the wave length.)

The singular values of a matrix.

$\lambda(M) = \lambda_0(M), \cdots \lambda_j(M), \cdots$ is the sequence of singular values of a matrix M, in decreasing order. When M is a symmetric positive semi-definite matrix (resp. operator), it is the spectrum of the matrix.

Let U be unitary. Then $\lambda(UA) = \lambda(AU) = \lambda(A)$.

The singular values of a matrix.

$\lambda(M) = \lambda_0(M), \cdots \lambda_j(M), \cdots$ is the sequence of singular values of a matrix M, in decreasing order. When M is a symmetric positive semi-definite matrix (resp. operator), it is the spectrum of the matrix.

Let U be unitary. Then $\lambda(UA) = \lambda(AU) = \lambda(A)$.

A fundamental inequality : Hoffman–Wielandt inequality for Hermitian matrices or operators, or for singular values

$$\sum (\lambda_j(C) - \lambda_j(D))^2 \leq \|C - D\|_{HS}^2.$$

Used to have approximation results in terms of the ℓ^2 norm.

Simplifications of the approximate model.

We start from the approximate model, given by the matrix

$$a^{\#}_{jk} = \frac{e^{2i\pi(h/\lambda + \frac{m}{2}(y_j - y'_k)^2)}}{h} = e^{+2i\pi(h/\lambda + \frac{m}{2}(y_j^2 + (y'_k)^2))} \times \frac{e^{-2i\pi m y_j y'_k}}{h}.$$

Multiplication of the jk entry by $e^{2i\pi h/\lambda} \times e^{i\pi m y_j^2} \times e^{i\pi m (y'_k)^2}$ corresponds to multiplication of the matrix, on the left and on the right, by diagonal matrices entries of modulus 1, that is, by unitary matrices. It does not change the singular values. So (if we use the previous approximation) we can replace a_{jk} by

$$\widetilde{a}_{jk} = \frac{\exp(-2i\pi m y_j y'_k)}{h}.$$

The mathematical question.

After approximation, simplification and normalization, the same authors reduce to the study of $n \times n$ matrices

$$A := \frac{\sqrt{m}}{n} \begin{pmatrix} e^{2i\pi m Z_1 Y_1} & \dots & e^{2i\pi m Z_1 Y_n} \\ \vdots & & \vdots \\ e^{2i\pi m Z_n Y_1} & \dots & e^{2i\pi m Z_n Y_n} \end{pmatrix}.$$

Here Y_j, Z_k independent and uniformly distributed in $I := (-1/2, +1/2)$.
So A is a random matrix, with alea given by the Z_j's and alea given by the Y_k's.

By assumption $m \ll n$ is large.
Describe the singular values of A (i.e. the spectrum of A^*A or AA^*) for m large, m/n small.
Typically, $m \approx n^\delta$, with $1/2 \leq \delta < 1$.

Finite Fourier Transform.

Kind of discretization of the operator on $L^2(I)$ given by

$$\mathcal{F}_m(f)(z) := \sqrt{m} \int_{-1/2}^{+1/2} e^{2i\pi mzy} f(y) dy.$$

Finite Fourier Transform.

Kind of discretization of the operator on $L^2(I)$ given by

$$\mathcal{F}_m(f)(z) := \sqrt{m} \int_{-1/2}^{+1/2} e^{2i\pi mzy} f(y)dy.$$

The compact operator $\mathcal{F}_m^* \mathcal{F}_m$ is given by

$$\mathcal{Q}_m(f)(x) = \mathcal{F}_m^* \mathcal{F}_m(f)(x) = \int_{-1/2}^{+1/2} \frac{\sin m\pi(x-y)}{\pi(x-y)} f(y)dy.$$

$\mathcal{Q}_m$ is the *time frequency limiting operator.*
$\frac{\sin \pi x}{\pi x}$ *is known as the sinc kernel.*

Eigenvalues of $\mathcal{Q}_m$ are ordered in a decreasing sequence, starting from 0.

Comparison Theorem for random matrices.

Theorem [BK]. The singular values of A are close to the singular values of $\mathcal{F}_m$. More precisely, for a given $\xi > 0$, with probability $1 - 2e^{-2\xi^2}$,

$$\left(\sum_{j \geq 0} [\lambda_j(A^*A) - \lambda_j(\mathcal{F}_m^* \mathcal{F}_m)]^2) \right)^{1/2} \leq \frac{4\xi m}{\sqrt{n}}.$$

$\sum_j \lambda_j(\mathcal{F}_m)^4 \approx m$. So the right hand side is an error if $\frac{m}{n}$ is small.
Probability 98% for $\xi \approx 1.6$. For $\xi = 2$, probability ≈ 0.9993.

Comparison Theorem for random matrices.

Theorem [BK]. The singular values of A are close to the singular values of $\mathcal{F}_m$. More precisely, for a given $\xi > 0$, with probability $1 - 2e^{-2\xi^2}$,

$$\left(\sum_{j\geq 0} [\lambda_j(A^*A) - \lambda_j(\mathcal{F}_m^*\mathcal{F}_m)]^2)\right)^{1/2} \leq \frac{4\xi m}{\sqrt{n}}.$$

$\sum_j \lambda_j(\mathcal{F}_m)^4 \approx m$. So the right hand side is an error if $\frac{m}{n}$ is small.
Probability 98% for $\xi \approx 1.6$. For $\xi = 2$, probability ≈ 0.9993.

Corollaries : good approximations (of the degrees of freedom, the capacity) are obtained by considering the finite Fourier transform instead of A.

Concentration inequalities.

The proof is divided into steps, first in the methods

- Estimates for the expectation.
- Use of concentration inequalities (McDiarmid's inequality, vector valued versions).

Secondly in the way to proceed

- Comparison of A^*A with its expectation in Z, given by

$$H := \mathbb{E}_Z(A^*A) = \frac{1}{n}\left(\frac{\sin(\pi m(Y_k - Y_j))}{\pi(Y_k - Y_j)}\right)_{j,k=1}^{n} := \frac{1}{n}\left(\kappa_m(Y_j, Y_k)\right)_{j,k=1}^{n}.$$

  $\kappa_m(x, y) = m\,\mathrm{sinc}(m(x-y))$ is a positive definite kernel.

- Comparaison of H with $\mathcal{Q}_m = \mathcal{F}_m^*\mathcal{F}_m$. Generalizes to all κ positive definite kernels.

A close question : kernel random matrices.

Assume more generally that $\mathcal{X}$ is a locally compact metric space and P a probability law on $\mathcal{X}$ here I and the Lebesgue measure. Let κ be a positive definite kernel on $\mathcal{X}$ here $m\,\mathrm{sinc}(mx)$.

Theorem. *Let κ be a positive definite kernel and T_κ be the integral operator with kernel κ. Assume that $m := \sup_y \kappa(y, y) < \infty$. For $(Y_k)_{k=1,\cdots n}$ i.i.d. with law P, let*

$$H_\kappa := n^{-1} \left(\kappa(Y_j, Y_k)\right)_{j,k=1,\cdots n}.$$

Then

$$\mathbb{E}\left(\sum_{j=0}^{n-1} |\lambda_j(H_\kappa) - \lambda_j(T_\kappa)|^2\right) \leq \frac{m^2}{n}.$$

Gives the error in $m/\sqrt{n}$.

A tool for big data.

Estimates for one eigenvalue are given by Shawe-Taylor, Cristianini and Kandola (see also Blanchard Bousquet and Zwald).
ℓ^2 estimates have been given by Adamczak et al., Rosasco et al.,...
Such a matrix is called a kernel matrix or a kernel Gram matrix.

For machine learning non linear principal component analysis (i. e. eigenvectors of such matrices) : one looks for the eigenfunctions and eigenvectors of a kernel matrix, for which scalar products are replaced by $\kappa(x, y)$, with κ a positive definite kernel.

Generalizations to other laws for the sample.

Assume that Z_j's and Y_j's are i.i.d. with law given by the symmetric probability μ on $\mathbb{R}$. Then we have the same kind of results, except that $\mathcal{F}_m$ is replaced by $\mathcal{F}_{m,\mu}$, given by

$$\mathcal{F}_{m,\mu} f(y) = \sqrt{m} \int_{\infty}^{\infty} e^{-2i\pi yz} f(z) d\mu(z).$$

The new kernel is $\widehat{\mu}(x-y)$.

Band limited functions and Slepian problem.

Let B_m be the Paley-Wiener space of functions with spectrum in $(-m/2, +m/2)$ and

$$\hat{f}(\xi) := \int_{-\infty}^{+\infty} e^{-2i\pi x\xi} f(x)dx.$$

its Fourier transform.

How large on $[-1/2, +1/2]$ is a function in B_m ?

- $$\lambda_0(m) = \max\left\{\int_{-1/2}^{+1/2} |f|^2 dx\,; \quad \|f\|_2 = 1, f \in B_m\right\}.$$

- $$\lambda_n(m) = \min_{f_1,\cdots f_n} \max\{\int_{-1/2}^{+1/2} |f|^2 dx\,;\; f \in V_n^{\perp}, \|f\|_2 = 1, f \in B_m\},$$

  where V_n is generated by $f_1, f_2, ..., f_n$.

Band and time limiting operator.

$$f(x) = \int_{-m/2}^{m/2} e^{2i\pi x\xi}\hat{f}(\xi)d\xi$$
$$= \int_{-1/2}^{1/2} e^{2i\pi mxy}g(y)dy = \mathcal{F}_m g(x)$$

for $|x| < 1/2$, with $g(y) = \sqrt{m}f(my)$.
Then $\lambda_n(m)$ are the eigenvalues of the compact operator $\mathcal{F}_m^*\mathcal{F}_m$

$$\mathcal{Q}_m(g)(x) = \mathcal{F}_m^*\mathcal{F}_m(g)(x) = \int_{-1/2}^{+1/} \frac{\sin m\pi(x-y)}{\pi(x-y)}g(y)dy.$$

Can also be interpreted as the product of an operator where one cuts the function on an interval, then its Fourier transform.

(Landau and Widom 1980) m tending to ∞ :
$N(\alpha) = \#\{\lambda_j(\mathcal{Q}_m);\ \lambda_j(\mathcal{Q}_m) > \alpha\}$ is such that

$$N(\alpha) = m + \left[\frac{1}{\pi^2}\log\left(\frac{1-\alpha}{\alpha}\right)\right]\log(m) + o(\log(m)).$$

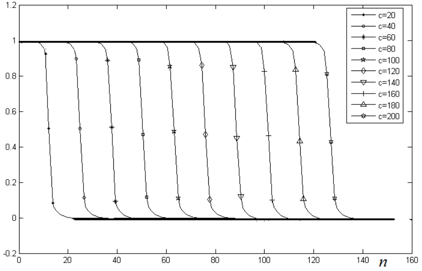


FIGURE: Graph of the $\lambda_j(\mathcal{Q}_m)$ for different values of $m = \frac{2c}{\pi}$

Non asymptotic results for $\lambda_j(\mathcal{Q}_m)$.

Landau (1991) : $\lambda_{\lceil m \rceil} \leq 1/2 \leq \lambda_{[m]-1}$.

A lucky incident (Slepian et al, Bell labs in the sixties) : The eigenfunctions $\psi_{s,m}$ (which are called Prolate Spheroidal Wave Functions) are also eigenfunctions of an explicit Sturm-Liouville operator on I

$$\mathcal{L}_m\psi := -\frac{d}{dx}\left[(1-4x^2)\psi'\right] + \pi^2 m^2 x^2 \psi.$$

Can be used to study the behavior of the eigenvalues of $\mathcal{Q}_m$. Many recent works (Osipov, Rokhlin-Xiao, Bonami-Karoui, ...).
In particular :

The right side of the plunge region

From now on, joint work with Jaming and Karoui.

Theorem. [B., Jaming, Karoui 2018] *there exists* a, C *such that, for* $s > m$,

$$\lambda_s(\mathcal{Q}_m) \leq Ce^{-a\left(\frac{s-m}{\log m}\right)}.$$

Can be compared with Landau-Widom estimate : if their asymptotic formula was exact, we would have

$$\lambda_s(\mathcal{Q}_m) \leq e^{-\pi^2\left(\frac{s-m}{\log m}\right)}.$$

It improves results of Osipov and Israel for $s > m$. But Israel gives also bounds for $s < m$.

$$\lambda_0(\mathcal{Q}_m) = \max \frac{\|\mathcal{F}_m f\|^2_{L^2(I)}}{\|f\|^2_{L^2(I)}}$$

$$\lambda_0(\mathcal{Q}_m) = \max \frac{\|\mathcal{F}_m f\|^2_{L^2(I)}}{\|f\|^2_{L^2(I)}} \geq 1 - 5e^{-\pi m/2}.$$

Proof. Test with $f(x) = \sqrt{2}m^{1/4}e^{-\pi m x^2}\chi_I(x) = g(x) - h(x)$, with g the normalized Gaussian function $g(x) = \sqrt{2}m^{1/4}e^{-\pi m x^2}$ and use the fact that $\sqrt{m}\int_{\mathbb{R}} e^{-2i\pi mxy} g(y)dy = g(x)$.

$$\lambda_0(\mathcal{Q}_m) = \max \frac{\|\mathcal{F}_m f\|^2_{L^2(I)}}{\|f\|^2_{L^2(I)}} \geq 1 - 5e^{-\pi m/2}.$$

Proof. Test with $f(x) = \sqrt{2}m^{1/4}e^{-\pi m x^2}\chi_I(x) = g(x) - h(x)$, with g the normalized Gaussian function $g(x) = \sqrt{2}m^{1/4}e^{-\pi m x^2}$ and use the fact that $\sqrt{m}\int_{\mathbb{R}} e^{-2i\pi mxy}g(y)dy = g(x)$.

Fuchs asymptotic result : $1 - \lambda_0(m) \simeq \frac{\pi\sqrt{m}}{\sqrt{2}}e^{-\pi m}$ for m tending to ∞.

Generalizes to eigenvalues up to $m/3$ by using Hermite functions.

$$\lambda_0(\mathcal{Q}_m) = \max \frac{\|\mathcal{F}_m f\|^2_{L^2(I)}}{\|f\|^2_{L^2(I)}} \geq 1 - 5e^{-\pi m/2}.$$

Proof. Test with $f(x) = \sqrt{2}m^{1/4}e^{-\pi m x^2}\chi_I(x) = g(x) - h(x)$, with g the normalized Gaussian function $g(x) = \sqrt{2}m^{1/4}e^{-\pi m x^2}$ and use the fact that $\sqrt{m}\int_{\mathbb{R}} e^{-2i\pi mxy}g(y)dy = g(x)$.

Fuchs asymptotic result : $1 - \lambda_0(m) \simeq \frac{\pi\sqrt{m}}{\sqrt{2}}e^{-\pi m}$ for m tending to ∞.

Generalizes to eigenvalues up to $m/3$ by using Hermite functions. As a consequence,

$$E(m) = \prod(1 - \lambda_k(\mathcal{Q}_m)) \leq e^{-0.15m^2}.$$

Can be compared with the asymptotic behaviour $E(m) \sim e^{-\pi^2 m^2/8+\cdots}$ given by Des Cloizeaux and Mehta in connection with the hole probability of GUE random matrices in the bulk.

Plunge region and Remez Inequality.

Which $n+1-$dimensional space to be used for $m \approx n$?

$$\lambda_n(\mathcal{Q}_m) = \max_{V_n} \min_{f \neq 0, f \in V_n} \frac{\|\mathcal{F}_m f\|^2_{L^2(I)}}{\|f\|^2_{L^2(I)}}.$$

Take for V_n the space of functions that may be written as $\sum_{j=0}^n a_j \varphi(x - \frac{j}{n+1} + \frac{1}{2})$, with φ compactly supported in $I/(n+1)$.

Plunge region and Remez Inequality.

Which $n+1-$dimensional space to be used for $m \approx n$?

$$\lambda_n(\mathcal{Q}_m) = \max_{V_n} \min_{f \neq 0, f \in V_n} \frac{\|\mathcal{F}_m f\|^2_{L^2(I)}}{\|f\|^2_{L^2(I)}}.$$

Take for V_n the space of functions that may be written as $\sum_{j=0}^n a_j \varphi(x - \frac{j}{n+1} + \frac{1}{2})$, with φ compactly supported in $I/(n+1)$.
To bound below the ratio, have to consider the Remez type best constant $\Gamma_2(n, \varepsilon)$ such that

$$\int_{-\varepsilon}^{+\varepsilon} |P(e^{it})|^2 \, dt \geq \Gamma(n, \varepsilon) \int_{-\pi}^{\pi} |P(e^{it})|^2 \, dt,$$

for any P is a Taylor polynomial of degree n and any J interval of length ε.

Plunge region and Remez Inequality.

Which $n+1-$dimensional space to be used for $m \approx n$?

$$\lambda_n(\mathcal{Q}_m) = \max_{V_n} \min_{f \neq 0, f \in V_n} \frac{\|\mathcal{F}_m f\|^2_{L^2(I)}}{\|f\|^2_{L^2(I)}}.$$

Take for V_n the space of functions that may be written as $\sum_{j=0}^n a_j \varphi(x - \frac{j}{n+1} + \frac{1}{2})$, with φ compactly supported in $I/(n+1)$.
To bound below the ratio, have to consider the Remez type best constant $\Gamma_2(n, \varepsilon)$ such that

$$\int_{-\varepsilon}^{+\varepsilon} |P(e^{it})|^2 \, \mathrm{d}t \geq \Gamma(n, \varepsilon) \int_{-\pi}^{\pi} |P(e^{it})|^2 \, \mathrm{d}t,$$

for any P is a Taylor polynomial of degree n and any J interval of length ε.

With $\varphi((n+1)x) = \mathbf{1}_{[-1/2,1/2]} \cos \pi x$, we find Landau's estimate $\lambda_{[m]-1} \geq 1/2$. Need only $\varepsilon = \pi$.