Spectral Asymptotics for Random DFT Submatrices

Brendan Farrell

Techische Universität Berlin
Berlin, Germany

École Supérieure d'Électricité
Paris, Dec. 17, 2008

Outline:

1. Uncertainty principles
2. Random approach
3. Results

The work here was largely motivated by questions raised in the work of Joel Tropp: "Linear Independence of Sines and Spikes", Journal of Fourier Analysis and Applications, 2008.

Our question originates with the discrete uncertainty principle:

We denote the Discrete Fourier Transform (DFT) matrix F, and we set $Fx = \hat{x}$, where $(Fx)_k = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} e^{-2\pi i (j-1)(k-1)/n}$.

Given two sets $T \subset \{1, ..., n\}$ and $\Omega \subset \{1, ..., n\}$, does there exist a vector $x \in \mathbb{C}^n$ such that $\text{supp}(x) \subset T$ and $\text{supp}(\hat{x}) \subset \Omega$?

Several well-known answers to this question:

Donoho and Stark, 1989: If $|T| \cdot |\Omega| < n$, then there does not exist an x satisfying $\mathrm{supp}(x) \subset T$ and $\mathrm{supp}(\hat{x}) \subset \Omega$.

Donoho and Stark, 1989: If $|T| + |\Omega| < 2\sqrt{n}$, then there is also x.

Tao, 2005: If n is prime and $|\Omega| + |T| \leq n$, then a corresponding x does not exist.

(We will come back to these later.)

Let's quantify the notion of existence. We can do this with the DFT matrix.

- Let $f_j(k) = \frac{1}{\sqrt{n}} e^{2\pi i (j-1)(k-1)/n}$, the j^{th} complex exponential. Unitary and self-adjoint.

- Then $F = [f_1|f_2|...|f_n]^*$

- Define the Restriction matrices R_T and R_Ω to have all entries off the diagonal equal to 0, and a 1 on the i, i entry of the diagonal if $i \in T$ or $i \in \Omega$.

- $R_T x = x$ or $R_\Omega \hat{x} = \hat{x}$ if and only if $\text{supp}(x) \subset T$ or $\text{supp}(\hat{x}) \subset \Omega$.

Suppose for a given pair T and Ω there exists x with $\text{supp}(x) \subset T$ and $\text{supp}(\hat{x}) \subset \Omega$.

Then $R_T x = x$, $\Rightarrow FR_T x = \hat{x}$, and $R_\Omega F R_T x = \hat{x} = Fx$.

So, if there exists an x corresponding to T, Ω, $\|R_\Omega F R_T\|_{\text{Op}} = 1$.

Suppose there does not exist such an x. Suppose, that $\text{supp}(x) \subset T$, but $\text{supp}(\hat{x}) \not\subset \Omega$.

Then

$$\|R_\Omega F R_T x\|_2 = \|R_\Omega F x\|_2 = \|R_\Omega \hat{x}\|_2 < \|\hat{x}\|_2 = 1,$$

where the inequality is strict, due to Plancharel's theorem.

A similar statement holds, of course, if the support is not contained in T or in neither T nor Ω.

Thus, $\|R_\Omega F R_T\|_{\text{Op}} = 1$ if and only if there exists a vector with supports T and Ω

We may thus address the existence question by looking at the norm of $R_\Omega F R_T$.

What does $R_\Omega F R_T$ look like?
Recall that $f_j(k) = \frac{1}{\sqrt{n}} e^{2\pi i (j-1)(k-1)/n}$ and $F = [f_1|f_2|\ldots|f_n]^*$.

Thus R_T leaves the columns corresponding to T (time) and R_Ω leaves the rows (frequencies) corresponding to Ω.

Existence of a vector x is then a consequence of norm of the DFT submatrix.

- We also define $F_{\Omega T}$ to be the matrix of dimension to be the restriction of F to the rows with indices in Ω and columns with indices in T.

- $F_{\Omega T}$ has dimension $\mathrm{Tr}(T) \times \mathrm{Tr}(\Omega)$ and removes the trivial zero eigenvalues from $R_\Omega F R_T$.

- While we are interested in the singular values of $F_{\Omega T}$, it is at first easier to work with $R_\Omega F R_T$.

We make the sets T and Ω random.

Definition

A square matrix is called a *diagonal projection matrix* if its off-diagonal entries all zero and its diagonal entries are zero or 1.

Definition

A random diagonal projection matrix will be called a *Bernoulli* diagonal projection matrix if the diagonal entries are independent and equal to 1 with probability p and 0 with probability $1 - p$.

A random diagonal projection matrix will be called a *uniform* diagonal projection matrix if the diagonal is given by a uniform distribution on the set of diagonals containing only 0's and 1's and having a specified trace.

Thus, rather than using the restriction matrices R_T and R_Ω, we use the diagonal projection matrices P, Q (different only in name), and investigate the spectral properties of PFQ.

The numbers p and q give the proportion of rows/columns that are set to zero.

That is, in expectation, PFQ will have $(1-p)n$ nonzero rows and $(1-q)n$ nonzero columns.

There are many existing results on the norm of the random matrix PFQ; however, to the best of my knowledge, not any asymptotic results.

As summarized and extended in [Tropp, 2008], previous non-asymptotic probabilistic results *on the largest eigenvalue* include:

- Candès and Romberg, 2006.
- Candès and Tao, 2006.
- Rudelson and Vershynin, 2006.
- Tropp, 2008.

Limiting empirical spectral distribution.

• We begin with $2n$ random variables, namely the diagonal elements of P and Q.

• From these we obtain n real random variables in the interval $[0,1]$, namely the eigenvalues of $P_nF_nQ_nF_n^*P_n$. We denote them $\{\lambda_i(P_nF_nQ_nF_n^*P_n)\}_{i=1}^n$.

• We are interested in the convergence of

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\delta_{\lambda_n(PFQF^*P)}(x)\xrightarrow{w.a.s.}\mu(x).$$

• Weak, almost-sure convergence of the empirical spectral distribution means, for any $f\in\mathcal{C}([0,1])$,

$$\lim_{n\to\infty}\langle f,\frac{1}{n}\sum_{i=1}^{n}\delta_{\lambda_n(PFQF^*P)}\rangle\xrightarrow{a.s.}\langle f,\mu(x)\rangle.$$

Recall that we are most interested in the largest/smallest eigenvalues of $F_{\Omega T}$.

We follow the strategy developed for the Wishart matrices:

1. Determine the limiting distribution of the bulk. (Done.)

2. Show that the probability of a single eigenvalue outside the limiting support decreases with dimension. (Conjectured.)

- Let H be an $n \times n$ matrix with $H_{i,j} \sim \mathcal{NC}(0, 1/\sqrt{n})$.

- Then HH^* is called a Wishart matrix.

Marčenko and Pastur used the Stieltjes transform to obtain the the limiting empirical spectral distribution for Wishart matrices[1] (Also Silverstein and Bai[2])

For a positive, real-valued random variable X, the Stieltjes transform $m_X(z)$ for $z \in \mathbb{C}^+$ (positive imaginary part) is

$$m_X(z) = \mathbb{E}(z - X)^{-1}.$$

In order to use this tool, we work with the (positive, real) eigenvalues of the matrix $PFQ(PFQ)^* = PFQF^*P$.

[1] Marčenko and Pastur: Distribution of eigenvalues of a class of random matrices. *Mat. USSR Sbornik*, 1:457-483, 1967.

[2] J.Silverstin and Z. Bai. On the empirical distribution of eigenvalues of a slcass of large dimensional random matrices. *Journal of Multivariate Analysis*, 55(2):175-192, 1995.

In order to determine $m_{X=\text{eig}(PFQF^*P)}(z)$, we start with the η-transform, introduced by Tulino and Verdu[3]:

$$\eta_X(z) = \mathbb{E}(1 + zX)^{-1}.$$

Once we have $\eta_X(z)$, we may recover $m_X(z)$ by using

$$m_X(z) = -\frac{1}{z}\sum_{k=0}^{\infty}(z)^{-k}\mathbb{E}[X^k] \tag{1}$$

and

$$\eta_X(z) = \sum_{k=0}^{\infty}(-z)^k\mathbb{E}[X^k], \tag{2}$$

so that

$$m_X(z) = -\frac{1}{z}\eta_X(-\frac{1}{z}). \tag{3}$$

---

[3] Tulino, Verdu: "Random Matrices and Wireless Communications", Foundations and Trends in Information Theory, 1(1), 2004.

Step 1: Let P_n, F_n and Q_n denote the matrices P, Q and F of dimension $n \times n$.

Lemma

If P and Q are Bernoulli projection matrices, then the η-transform of $P_n F_n Q_n F_n^* P_n$ converges almost surely to the solution of the fixed-point equation

$$\eta_{PFQF^*P}(z) = \eta_Q\left(z - z\frac{e}{\eta_{PFQF^*P}(z)}\right). \tag{4}$$

The idea for this lemma comes from the work of Tulino, Verdu, Caire and Shamai[4]. A central piece of the proof relies on work developed in a preprint of Tulino et al. for their similar setting.

[4] Capacity of the gaussian erasure channel, ISIT, June 2007.

Step 2: we solve the previous equation for $\eta_{PFQF^*P}(z)$ and determine $m_{PFQF^*P}(z)$.

Lemma

Let P_n and Q_n be independent Bernoulli with (expected) traces $1-p \in [0,1]$ and $1-q \in [0,1]$ respectively. Then the η-transform of $P_n F_n Q_n F_n^* P_n$ converges almost surely to the asymptotic η-transform

$$\eta_{PFQF^*P}(z) =$$

$$\frac{1+(p+q)z+\sqrt{1+(2(p+q)-4pq)z+((p+q)^2-4pq)z^2}}{2(1+z)}.$$

We are now in a situation very similar to the Marčenko-Pastur distribution[5].

Step 3: Stieltjes inversion formula. (See pp 92-94 of Hiai and Petz: The Semicircle Law.... for good explanation.)

If/where the random variable X has a density, it can be recovered by the Stieltjes inversion formula[6]

$$\frac{dF_X(x)}{dx} = \frac{1}{\pi} \lim_{\omega \to 0} \text{Im } m_X(x + i\omega). \tag{5}$$

We may apply this tool for $x \in (0, 1)$, but not for $x = 0, 1$.

[5] Marčenko and Pastur: Distribution of eigenvalues of a class of random matrices. *Mat. USSR Sbornik*, 1:457-483, 1967.

[6] See, for example, Silverstein and Bai, 1995

Step 4: the measures at 0 and 1.

Clearly there is a point mass at 0, since $\operatorname{rank}(PFQF^*P) \leq \operatorname{rank}(P)$.

We determine $\mu(0)$ by determining $\lim_{r\to\infty} \eta_{PFQF^*P}(r)$.

We determine $\mu(1)$ by working with

$$\eta_{PFQF^*P}(z) = \sum_{k=0}^{\infty}(-z)^k \int_0^1 x^k \frac{dF(x)}{dx} + \frac{\mu(1)}{1+z}.$$

Theorem

For $i = 1, ..., n$ let i be contained in Ω_n independently with probability $(1-q)$ and, also independently, let i be included in T_n with probability $(1-p)$. Then the empirical eigenvalue distribution of the $\min(|T_n|, |\Omega_n|)$ largest eigenvalues of $F_{\Omega_n T_n} F^*_{\Omega_n T_n}$ converges almost surely to

$$\frac{dF_{F_{q,p}F^*_{q,p}}(x)}{dx} = \frac{\sqrt{(1-\frac{r_-}{x})(\frac{r_+}{x}-1)}}{2\pi(1-x)(1-\max(p,q))} \cdot I_{(r_-,r_+)}(x) + \frac{\max(0, 1-(p+q))}{1-\max(p,q)} \cdot \delta(x-1) \tag{6}$$

where

$$r_- = (\sqrt{p(1-q)} - \sqrt{q(1-p)})^2 \tag{7}$$

and

$$r_+ = (\sqrt{p(1-q)} + \sqrt{q(1-p)})^2. \tag{8}$$

Corollary

For $i = 1, \ldots, n$ let i be contained in Ω_n independently with probability $(1 - q)$ and, also independently, let i be included in T_n with probability $(1 - p)$. Then the empirical singular values distribution of $F_{\Omega_n T_n}$ converges almost surely to

$$\frac{dF^s_{F_{q,p}}(x)}{dx} = \frac{\sqrt{x^2(1 - \frac{r_-}{x})(\frac{r_+}{x} - 1)}}{2\pi(1 - x^2)(1 - \max(p, q))} \cdot I_{(r_-, r_+)}(x) + \frac{\max(0, 1 - (p + q))}{1 - \max(p, q)} \cdot \delta(x - 1) \tag{9}$$

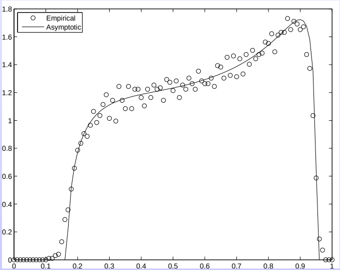
where

$$r_- = (\sqrt{p(1 - q)} - \sqrt{q(1 - p)})^2 \tag{10}$$

and

$$r_+ = (\sqrt{p(1 - q)} + \sqrt{q(1 - p)})^2. \tag{11}$$

Here we plot the empirical distribution. The original matrix size is 500×500, and we create 100 such matrices.



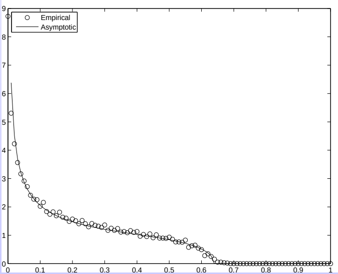
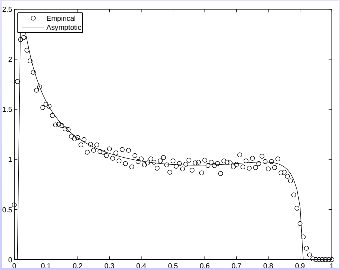


Figure: left:p=0.4,q=0.8. right: p=0.8,q=0.8.

Again, the original matrix size is 500×500, and we create 100 such matrices.



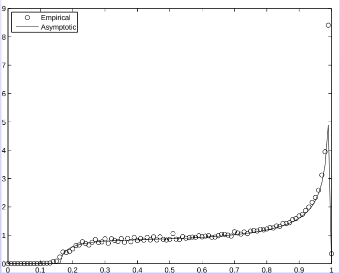


Figure: left:p=0.6,q=0.8. right: p=0.6,q=0.3.

We have shown

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_n(PFQF^*P)}(x) \xrightarrow{w.a.s.} \mu(x),$$

where $\mathrm{supp}(\mu) \subset [r_-, r_+]$.
However, we can change one value and still have the same *w.a.s.* convergence:

$$\lim_{n\to\infty} \left\{ \frac{1}{n} \sum_{i=1}^{n-1} \delta_{\lambda_n(PFQF^*P)}(x) + \frac{1}{n} \delta_{\left(\frac{r_+ + 1}{2}\right)}(x) \right\} \xrightarrow{w.a.s.} \mu(x).$$

Thus, we cannot yet make any statement about the largest eigenvalue, i.e. the norm!

In the case of Wishart matrices, Bai and Silverstein[7] and Bai, Silverstein and Yin[8] showed that the probability of *an* eigenvalue outside the limiting support converges to zero.

However, in the Wigner case, the largest eigenvalue does not converge to a value, but to the Tracy-Widom distribution.

It is unclear in which of these two situations our case falls. (Hopefully the former.)

[7] Z.D. Bai and J.W. Silverstein: No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices. *Ann. Probab.* 26(1):316-345.

[8] Z.D. Bai, J.W. Silverstein and Y.Q. Yin: A note on the largest eigenvalue of of a large dimensional sample covariance matrix. *J. Multiv. Anal.* 26:166-168

By looking at the square root of the eigenvalues of $PFQF^*P$, we can answer a small part of our original question.

Corollary

For $n \in \mathbb{N}$, let Ω_n and T_n be random subsets of $\{1, \ldots, n\}$ given according to a Bernoulli distribution on $\{1, \ldots, n\}$ with probability of inclusion $1 - p$ and $1 - q$. If $p + q \leq 1$, then $\|F_{\Omega_n T_n}\|$ converges almost surely to 1.

We would like to make a statement about the largest eigenvalue when $p + q > 1$.

As a special case of the previous result, we prove one part of Conjecture 18 from [Tropp, 2008].

Corollary

For $n \in \mathbb{N}$, let Ω_n and T_n be random subsets of $\{1, \ldots, n\}$, both given according to a Bernoulli distribution on $\{1, \ldots, n\}$ with probability of inclusion $1 - p$. If $p \leq 1/2$, then $\|F_{\Omega_n T_n}\|$ converges almost surely to 1.

Again, we still don"t have a statement for $p > 1/2$.

Thank you!