Random matrix theory: lecture 10

Asymptotic analysis of deterministic (Toeplitz) matrices

[ref: Bob Gray's report]

Circulant matrices

\[
C = \begin{pmatrix}
  c_0 & c_1 & \cdots & c_{n-1} \\
  c_1 & c_0 & \cdots & c_{n-2} \\
  \vdots & \vdots & \ddots & \vdots \\
  c_{n-1} & c_{n-2} & \cdots & c_0
\end{pmatrix}
\]

cyclic shifts to the right

nxn matrix, \( c_0 \ldots c_{n-1} \in \mathbb{C} \)

notation: \( C = \text{circ} (c_0, c_1, \ldots, c_{n-1}) \)

Lemma

Let \( \alpha \in \mathbb{C} \) be such that \( \alpha^n = 1 \) \((n^{th} \text{ root of unity})\).

Then \( \mathbf{u} = \left( \alpha^0, \alpha^1, \ldots, \alpha^{n-1} \right) \) is an eigenvector of \( C \)

with corresponding eigenvalue \( \lambda = \sum_{e=0}^{n-1} c_e \alpha^e \)

Proof

One has to check that \( Cu = \lambda u \):

\[
(Cu)_j = \sum_{k=1}^{n} C_{jk} u_k = \sum_{k=1}^{j} c_{n-j+k} \alpha^k + \sum_{k=j}^{n} c_{k+j} \alpha^k
\]

\[
= \sum_{l=n-j+1}^{n-1} c_l \alpha^{l+j-n} + \sum_{e=0}^{n-1} c_e \alpha^{l+ij}
\]

\[
= \alpha^{l+ij} \text{ since } c_{n-j} = 1
\]

\[
= \sum_{e=0}^{n-1} c_e \alpha^{l+ij} = \left( \sum_{e=0}^{n-1} c_e \alpha^e \right) \alpha^j = \lambda u_j
\]
There are $n$ different $\alpha$'s such that $\alpha^n = 1$:

$$\alpha_k = \exp\left(\frac{2\pi i k}{n}\right) \quad k=1 \ldots n$$

and it turns out that the $n$ different eigenvectors $u_1 \ldots u_n$ generated by $\alpha_1 \ldots \alpha_n$ are orthogonal.

$\Rightarrow$ Proposition: There exist $U$ unitary and $\Lambda = \text{diag}(\lambda_1 \ldots \lambda_n)$ such that $C = U \Lambda U^*$, where

$$\begin{align*}
\lambda_k &= \sum_{e=0}^{n-1} C e (\alpha_k^e) e = \sum_{e=0}^{n-1} C e \exp\left(\frac{2\pi i e k}{n}\right) \\
u_{jk} &= \frac{1}{n} (\alpha_k^j) = \frac{1}{n} \exp\left(\frac{2\pi i j k}{n}\right) \quad \text{[DFT matrix]}
\end{align*}$$

Consequences

- All circulant matrices share the same set of eigenvectors!

Only the eigenvalues depend (linearly!) on the values of $\alpha_1 \ldots \alpha_n$!

- If $C$ is a circulant matrix, then:
  - $C^*$ is circulant: $CC^* = C^*C$, i.e. $C$ is normal
  - If $C$ is invertible, then $C^{-1}$ is circulant

- If $C_1, C_2$ are circulant with eigenvalues $\lambda_1, \lambda_2$, and $\alpha, \beta \in \mathbb{C}$:
  - $\alpha C_1 + \beta C_2$ is circulant, with eigenvalues $\lambda_k = \alpha \lambda_1 + \beta \lambda_2$
  - $C_1 C_2$ is circulant, with eigenvalues $\lambda_k = \lambda_1 \lambda_2$
(Finite-order) Toeplitz matrices

\[
T^{(n)} = \begin{pmatrix}
    t_0 & t_1 & \cdots & t_k & 0 \\
    t_1 & t_0 & & t_k & 0 \\
    \vdots & \vdots & \ddots & \vdots & \vdots \\
    t_k & 0 & \cdots & t_0 & t_1 \\
    0 & t_k & \cdots & 0 & t_0 \\
\end{pmatrix}
\]

\(n \times n\) matrix, \((T^{(n)})_{jk} = t_{|k-j|} \in \mathbb{C}\)

and \(t_0 = 0\) if \(k > 0\).

Let \(\lambda_1^{(n)} \ldots \lambda_n^{(n)}\) be the eigenvalues of \(T^{(n)}\); we are interested
in the asymptotic behaviour of these eigenvalues as \(n \to \infty\).

First remark: contrary to circulant matrices, there is no general
expression at finite \(n\) for the eigenvalues of \(T^{(n)}\) in terms
of the numbers \(t_k\); and the DFT matrix is not the matrix
of eigenvectors of \(T^{(n)}\).

Let us define \(q(x) = \sum_{k=0}^{\infty} \frac{t_k}{e_0} e^{i\alpha_k} x\), \(x \in [0, 2\pi]\).

\(q\) is complex-valued, bounded and continuous,
\(t_k\) are the Fourier coefficients of \(q\): \(t_k = \frac{1}{2\pi} \int_0^{2\pi} q(x) e^{-i\alpha x} dx\).

Lemma 1 (connection between the \( \lambda's \) and \(q\))

For any \(m \geq 0\), we have

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (\lambda_k^{(n)})^m = \frac{1}{2\pi} \int_0^{2\pi} (q(x))^m dx
\]
Proof idea (details are left to homework 🙂)

- Consider the case \( b_0 = 1 \) for simplicity; the matrices

\[
\mathbf{T}^{(n)} = \begin{pmatrix} t_0 & t_1 & \cdots & 0 \\ 0 & t_1 & \cdots & t_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{n-1} \end{pmatrix}
\quad \text{and} \quad
\mathbf{C}^{(n)} = \begin{pmatrix} t_0 & t_1 & \cdots & 0 \\ 0 & t_1 & \cdots & t_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{n-1} \end{pmatrix}
\]

with e.v. \( \mathbf{A}_k^{(n)} \) and e.v. \( \mathbf{I}_k^{(n)} \)

can be shown to be "asymptotically equivalent,”

which implies that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (A_k^{(n)})^m = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (I_k^{(n)})^m
\]

\( \forall m \geq 0 \)

- Note that \( \mathbf{C}^{(n)} \) is circulant, so

\[
\mathbf{M}_k^{(n)} = \sum_{c=-c_0}^{c_0} e^{i2\pi kc/n} = g\left(\frac{2\pi k}{n}\right)
\]

(for \( n \geq 2b_0+1 \))

- Therefore,

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (A_k^{(n)})^m = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} g\left(\frac{2\pi k}{n}\right)
\]

(Riemann sums)

\[
= \frac{1}{2\pi} \int_0^{2\pi} g(x) \, dx
\]

Illustration

for \( m = 1 \):

\[
\sum_{k=1}^{n} (A_k^{(n)})^m \rightarrow \int_0^{2\pi} g(x) \, dx
\]

NB: The above result actually says that

\[
\lim_{n \to \infty} \frac{1}{n} \text{Tr} \left( (T^{(n)})^m \right) = \frac{1}{2\pi} \int_0^{2\pi} (g(x))^m \, dx \quad \forall m \geq 1
\]
Assumption $H$: $T_e = F_e$ for all $1 \leq i \leq n$.

Under assumption $H$, $T^{(n)}$ is Hermitian, so $T_i^{(n)}, i = 1, \ldots, n \in \mathbb{R}^n, \forall n$.

- $g$ is real-valued.

Lemma 2 (proof → homework again :)

Under assumption $H$, $a \leq T_i^{(n)} \leq b$ \quad $1 \leq k \leq n$

where $a := \inf_{x \in [0, b]} g(x) \leq \sup_{x \in [0, a]} g(x) =: b$.

Theorem (Grenander-Szegö 1958, Gray 1972)

Under assumption $H$, we have for any continuous function $f: [a, b] \to \mathbb{R}$:

$$
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(T_i^{(n)}) = \frac{1}{b-a} \int_{[a, b]} f(g(x)) \, dx. \quad \varepsilon \in [a, b] \text{ by lemma 2.} \quad \varepsilon \in [a, b] \text{ by def.}
$$

Proof:

- Lemma 1 proves the theorem for $f(y) = y^m$, $y \in [a, b]$.
- By linearity of the sum & integral, the theorem also holds for any $f$ of the form $f(y) = \sum_{m=0}^{m_0} c_m y^m$, i.e., any polynomial.
- By Weierstrass theorem, any continuous function $f$ on $[a, b]$ may be approximated uniformly by a sequence of polynomials, so the theorem extends to continuous functions.
Example: let $t_0 = 2$, $t_1 = t_2 = -1$, $t_3 = 0$ for all $l > 1$

\[
T^{(n)} = \begin{pmatrix}
2 & -1 & 0 \\
-1 & 0 & 1 \\
0 & 1 & 2
\end{pmatrix}, \quad q(x) = 2 - e^{ix} - e^{-ix} = 2(1 - \cos x)
\]

i.e. $a = 0$, $b = 1$

\[
g(\omega) = \frac{\sin \omega}{\omega}
\]

Important remark

Without assumption $H$, the theorem fails!

Counter-example: let $t_0 = 1$, $t_1 = -1$, $t_3 = 0$ otherwise

\[
T^{(n)} = \begin{pmatrix}
1 & -1 & 0 \\
-1 & 0 & 1 \\
0 & 1 & 1
\end{pmatrix}, \quad q(x) = 1 - e^{ix} \in \mathbb{C},
\]

but all eigenvalues of $T^{(n)}$ are equal to 1!

It still holds that $\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (A_k)^m = 1 = \frac{1}{2\pi} \int_0^{2\pi} (1 - e^{ix})^m \, dx$

but $\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} f(A_k) = f(1) = \frac{1}{2\pi} \int_0^{2\pi} f(1 - e^{ix}) \, dx$

for any continuous $f : \mathbb{C} \to \mathbb{C}$

Take home message:

When dealing with sequences of non-Hermitian matrices $A^{(n)}$, knowing $\lim_{n \to \infty} \frac{1}{n} \text{Tr}((A^{(n)})^m)$ $\forall m > 0$

is not sufficient to determine the asymptotic behaviour of eigenvalues.
Another perspective on the Grenander-Szegö theorem

- By the change of variable \( y = g(x) \), we have

\[
\frac{1}{va} \int_a^b f(g(x)) \, dx = \int_a^b f(y) \, p(y) \, dy \quad \text{for some } p(y)
\]

[NB: This works only if \( g \) is (piecewise) 1-to-1]

Choosing \( f(y) = 1 \), we get

\[
\frac{1}{va} \int_a^b 1 \, dx = 1 = \int_a^b p(y) \, dy
\]

and \( \int_a^b f(y) p(y) \, dy \geq 0 \) for any \( f(y) \geq 0 \), so \( p(y) \geq 0 \),

i.e. \( p(y) \) is the density of a (probability) distribution \( \mu \) on \( \mathbb{R} \).

Notation: \( \int_a^b f(y) \, p(y) \, dy = \int_a^b f(y) \, d\mu(y) \)

\[
\frac{d\mu}{dy} = p(y)
\]

- Recall now Dirac's \( \delta \)-distribution on \( \mathbb{R} \):

for \( c \in \mathbb{R} \), \( \delta_c \) is the distribution s.t. \( \int_{\mathbb{R}} f(y) \, d\delta_c(y) = f(c) \quad \forall f \)

The empirical eigenvalue distribution of \( T_n \) is defined as:

\[
\mu_n := \frac{1}{n} \sum_{k=1}^{n} \delta_{\lambda_k^{(n)}} \quad \text{(supported on \([a,b]\) by lemma 2)}
\]

i.e. \( \int_{\mathbb{R}} f(y) \, d\mu_n(y) = \frac{1}{n} \sum_{k=1}^{n} f(\lambda_k^{(n)}) \quad \forall f : [a,b] \to \mathbb{R} \)

- What Grenander-Szegö's thm. says is therefore:

\[
\lim_{n \to \infty} \int_{\mathbb{R}} f(x) \, d\mu_n(x) = \int_{\mathbb{R}} f(y) \, d\mu(y) \quad \forall f : [a,b] \to \mathbb{R} \text{ continuous}
\]

i.e. the sequence of distributions \( (\mu_n)_{n \geq 1} \) converges

weakly to the distribution \( \mu \) as \( n \to \infty \).
Example: \( t_0 = 2, t_1 = t_2 = -1, t_e = 0 \) for \( l > 1 \)

\[ q(x) = 2(1 - \cos x) = 4 \sin^2 \left( \frac{x}{2} \right) \]

Limiting eigenvalue profile

\[ \frac{1}{n} \sum_{k=1}^{n} f_k^{(n)} \to \frac{1}{2a} \int_{0}^{a} f \left( 4 \sin^2 \left( \frac{x}{2} \right) \right) dx \]

\[ = \frac{1}{a} \int_{0}^{a} f \left( 4 \sin^2 \left( \frac{x}{2} \right) \right) dx \]

\[ dy = 4 \sin \left( \frac{x}{2} \right) \cos \left( \frac{x}{2} \right) \frac{1}{2} dx \]

\[ \Rightarrow dx = \frac{2}{\sqrt{y(4-y)}} dy, \quad x=0 \leftrightarrow y=0, \quad x=a \leftrightarrow y=4 \]

\[ \frac{1}{n} \sum_{k=1}^{n} f_k^{(n)} \to \int_{0}^{a} f(y) \frac{2}{\pi \sqrt{y(4-y)}} dy \]

\[ p(y) = \frac{2}{\pi \sqrt{y(4-y)}} \cdot \mathbb{1}_{[0,4]}(y) \]

Limiting eigenvalue distribution

\[ \text{Limiting eigenvalue profile} \quad \text{Limiting eigenvalue distribution} \]

\[ \text{NB: it is therefore possible to talk about "eigenvalue distribution" even for deterministic matrices!} \]
Further generalizations of the theorem

1. If the sequence \( (t_e, e \in \mathbb{Z}) \) is not of finite order but satisfies still \( \sum_{e \in \mathbb{Z}} |t_e| < \infty \), then a proof similar to the preceding leads to the same result (but the approximation of \( T^{(n)} \) by a circulant matrix is more involved, since every entry of \( T^{(n)} \) is possibly non-zero).

ref: Gray's web report

2. If only the weaker condition \( \sum_{e \in \mathbb{Z}} |t_e^2| < \infty \) is satisfied, then the proof gets even more involved, but the result still holds true (ref: Grenander, Szegö).

3. The following is not exactly a generalization, but rather a rephrasing of the theorem:

Let \( f(y) = 1_{y \leq t} \) for some fixed \( t \in [a, b] \).

\( f \) is not continuous on \([a, b]\), but can be approximated by continuous functions on \([a, b]\), so it can be shown that the theorem still holds true for such \( f \), i.e. that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} 1_{x_{e^k} \leq t} = \frac{1}{2a} \int_{0}^{2\pi} f(x) e^{it} \, dx \quad \forall t \in [a, b]
\]
Let us define \( F_n(t) := \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\lambda_k^{(n)} \leq t} \) 
and \( F(t) := \frac{1}{2a} \int_{-a}^{a} \mathbb{1}_{g(x) \leq t} \, dx \). \( t \in [a, b] \)

Note that
- \( F_n(t) = \frac{1}{n} \# \left\{ k : \lambda_k^{(n)} \leq t \right\} \) is the proportion of eigenvalues of \( T^{(n)} \) less than or equal to \( t \).
- By the same change of variable as above \( (y = g(x)) \), \( F(t) = \int_{a}^{b} p(y) \, dy \).

Both \( F_n \) and \( F \) are therefore cumulative distribution functions, and the theorem says that

\[ F_n(t) \xrightarrow{\text{n \to \infty}} F(t) \quad \forall t \in [a, b] \]

which is a second characterization of the weak convergence of the corresponding sequence of distributions.

\((\ast)\) i.e. \( F_n(a) = 0, F_n(b) = 1, F_n \) is non-decreasing on \([a, b]\) 

(and \( F_n \) is a right-continuous function on \([a, b]\))