Random matrix theory: lecture 10

Asymptotic analysis of deterministic (Toeplitz) matrices

[ref: Bob Gray's report]

Circulant matrices

$$C = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ & & \ddots & \\ c_1 & \cdots & c_{n-1} & c_0 \end{pmatrix} \downarrow \text{cyclic shifts to the right}$$

$n \times n$ matrix, $c_0 \cdots c_{n-1} \in \mathbb{C}$

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

Lemma

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

Then $u = \begin{pmatrix} \alpha \\ \alpha^2 \\ \vdots \\ \alpha^{n-1} \\ 1 \end{pmatrix}$ is an eigenvector of C

with corresponding eigenvalue $\lambda = \sum_{\ell=0}^{n-1} c_\ell \alpha^\ell$

Proof

One has to check that $Cu = \lambda u$:

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

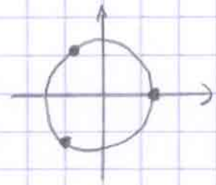
$\ell = n-j+k$ $\qquad$ $\ell = k-j$

$$= \sum_{\ell=n-j+1}^{n-1} c_\ell \underbrace{\alpha^{\ell+j-n}}_{=\alpha^{\ell+j} \text{ since } \alpha^{-n}=1} + \sum_{\ell=0}^{n-j} c_\ell \alpha^{\ell+j}$$

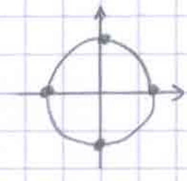
$$= \sum_{\ell=0}^{n-1} c_\ell \alpha^{\ell+j} = \left(\sum_{\ell=0}^{n-1} c_\ell \alpha^\ell\right) \alpha^j = \lambda u_j \quad \#$$

There are n different α's such that $\alpha^n = 1$:

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



$n=3$



$n=4$

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

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

$$\begin{cases} \lambda_k = \sum_{\ell=0}^{n-1} c_\ell (\alpha_k)^\ell = \sum_{\ell=0}^{n-1} c_\ell \exp\left(\frac{2\pi i k \ell}{n}\right) \\ U_{jk} = \frac{1}{\sqrt{n}} (\alpha_k)^j = \frac{1}{\sqrt{n}} \exp\left(\frac{2\pi i j k}{n}\right) \quad \text{[DFT matrix]} \end{cases}$$

Consequences

- All circulant matrices share the same set of eigenvectors! Only the eigenvalues depend (linearly!) on the values of $c_1..c_n$!
- 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 $\lambda_k^{(1)}, \lambda_k^{(2)}$, and $\alpha, \beta \in \mathbb{C}$:
  - $\alpha C_1 + \beta C_2$ is circulant, with eigenvalues $\lambda_k = \alpha \lambda_k^{(1)} + \beta \lambda_k^{(2)}$
  - $C_1 C_2$ is circulant, with eigenvalues $\lambda_k = \lambda_k^{(1)} \lambda_k^{(2)}$

(Finite-order) Toeplitz matrices

$$T^{(n)} = \begin{pmatrix} t_0 & t_1 & \cdots & t_{\ell_0} & & 0 \\ t_{-1} & \ddots & & & \ddots & t_{\ell_0} \\ t_{-\ell_0} & & \ddots & & & \vdots \\ & \ddots & & & \ddots & t_1 \\ 0 & & t_{-\ell_0} & \cdots & t_{-1} & t_0 \end{pmatrix} \quad n\times n \text{ matrix}, \quad \left(T^{(n)}\right)_{jk} = t_{k-j} \in \mathbb{C}$$

and $t_\ell = 0$ if $|\ell| > \ell_0$.

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_ℓ; and the DFT matrix is not the matrix of eigenvectors of $T^{(n)}$.

Let us define $g(x) := \sum_{\ell=-\ell_0}^{\ell_0} t_\ell e^{i\ell x}$, $x \in [0, 2\pi]$

g is complex-valued, bounded and continuous,

t_ℓ are the Fourier coefficients of g: $t_\ell = \frac{1}{2\pi} \int_0^{2\pi} g(x) e^{-i\ell x} dx$

Lemma 1 (connection between the λ's and g)

For any $m \geq 0$, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \left(\lambda_k^{(n)}\right)^m = \frac{1}{2\pi} \int_0^{2\pi} \left(g(x)\right)^m dx$$

18.4.07

4

<u>Proof idea</u> (details are left to homework :))

- consider the case $\ell_0 = 1$ for simplicity; the matrices

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

with e.v. $\lambda_k^{(n)}$ with e.v. $\mu_k^{(n)}$

can be shown to be "asymptotically equivalent", which implies that $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\lambda_k^{(n)})^m = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\mu_k^{(n)})^m$ $\forall m \geq 0$

- note that $C^{(n)}$ is circulant, so

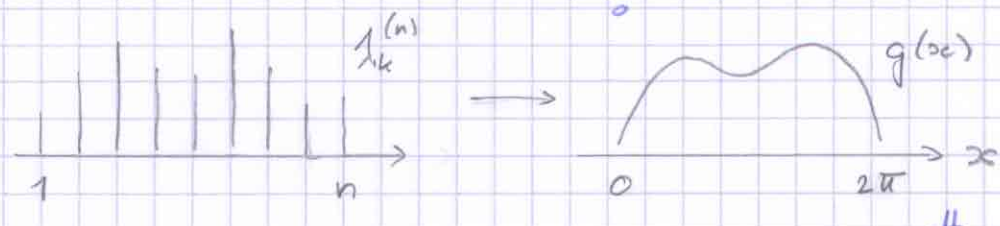
$$\mu_k^{(n)} = \sum_{\ell=-\ell_0}^{\ell_0} t_\ell \exp\left(\frac{2\pi i k \ell}{n}\right) = g\left(\frac{2\pi k}{n}\right)$$

(for $n \geq 2\ell_0+1$)

- Therefore, $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\lambda_k^{(n)})^m = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} g\left(\frac{2\pi k}{n}\right)$

(Riemann sums) $= \frac{1}{2\pi} \int_0^{2\pi} (g(x))^m \, dx$

illustration for m=1 :



#

<u>NB</u>: The above result actually says that

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{Tr}\left(\left(T^{(n)}\right)^m\right) = \frac{1}{2\pi} \int_0^{2\pi} (g(x))^m \, dx \quad \forall m \geq 1$$

Assumption H : $t_{-\ell} = \overline{t_\ell}$ for all $|\ell| \le \ell_0$

× Under assumption H: • $T^{(n)}$ is Hermitian, so $\lambda_1^{(n)} .. \lambda_n^{(n)} \in \mathbb{R}$, $\forall n$.

× • g is real-valued

Lemma 2 (proof → homework again ∵)

Under assumption H, $a \le \lambda_k^{(n)} \le b \quad \forall 1 \le k \le n$

where $a := \inf_{x\in[0,2\pi]} g(x) \le \sup_{x\in[0,2\pi]} g(x) =: b$

Theorem (Grenander - Szegö 1958, Gray 1972)

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

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\underbrace{\lambda_k^{(n)}}_{\in [a,b] \text{ by lemma 2}}) = \frac{1}{2\pi} \int_0^{2\pi} f(\underbrace{g(x)}_{\in [a,b] \text{ by def.}})\, dx$$

Proof

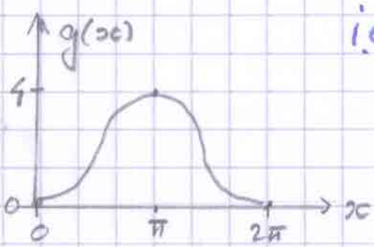
- lemma 1 proves the thm for f of the form $f(y) = y^m$, $y \in [a,b]$.
- by linearity of the sum & integral, the theorem also holds for any f of the form $f(y) = \sum_{m=0}^{m_0} c_m y^m$, i.e. any polynomial.
- × by Weierstrass theorem, any continuous function f on $[a,b]$ may be approximated uniformly by a sequence of polynomials, so the thm extends to continuous functions. #

Example: let $t_0 = 2$, $t_1 = t_{-1} = -1$, $t_\ell = 0 \quad \forall |\ell| > 1$

$$T^{(n)} = \begin{pmatrix} 2 & -1 & & 0 \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{pmatrix}, \quad g(x) = 2 - e^{ix} - e^{-ix} = 2(1-\cos x)$$

i.e. $a = 0$, $b = 4$

limiting eigenvalue profile



Important remark

Without assumption H, the theorem fails!

(Counter)-example: let $t_0 = 1$, $t_1 = -1$, $t_\ell = 0$ otherwise

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

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

it still holds that $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\lambda_k)^m = 1 = \frac{1}{2\pi} \int_0^{2\pi} (1-e^{ix})^m dx \quad \forall m \geq 0$

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

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

Take home message:

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

$$\frac{1}{2\pi}\int_0^{2\pi} f(g(x))\,dx = \int_a^b f(y)\,p(y)\,dy \quad \text{for some } p(y)$$

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

Choosing $f(y) \equiv 1$, we get $\frac{1}{2\pi}\int_0^{2\pi} 1\,dx = 1 = \int_a^b p(y)\,dy$

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 μ on $\mathbb{R}$.

notation: $\int_a^b f(y)\,p(y)\,dy = \int_a^b f(y)\,d\mu(y)$ $\quad \left(\frac{d\mu}{dy} = p(y)\right)$

- Recall now Dirac's δ-distribution on $\mathbb{R}$:

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

The empirical eigenvalue distribution of $T^{(n)}$ is defined as:

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

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

- What Grenanger-Szegö's thm says is therefore:

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

i.e. the sequence of distributions $(\mu_n)_{n\geq 1}$ converges weakly to the distribution μ as $n \to \infty$.

<u>Example</u>: $t_0 = 2,\ t_1 = t_{-1} = -1,\ t_\ell = 0 \quad \forall |\ell| > 1$

$\Rightarrow g(x) = 2(1-\cos x) = 4\sin\left(\frac{x}{2}\right)^2$ limiting eigenvalue profile

$$\frac{1}{n}\sum_{k=1}^{n} f(\lambda_k^{(n)}) \xrightarrow[n\to\infty]{} \frac{1}{2\pi}\int_0^{2\pi} f\left(4\sin\left(\frac{x}{2}\right)^2\right) dx$$

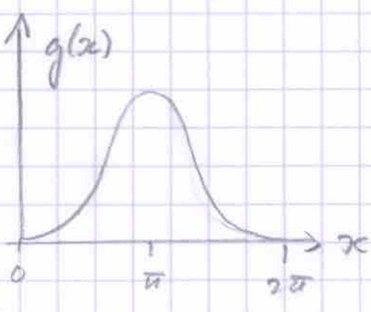
$$= \frac{1}{\pi}\int_0^{\pi} f\Big(\underbrace{4\sin\left(\tfrac{x}{2}\right)^2}_{=y}\Big)\, dx$$

$$dy = 4\sin\left(\frac{x}{2}\right)\cos\left(\frac{x}{2}\right)\frac{1}{2}\, dx\,.$$

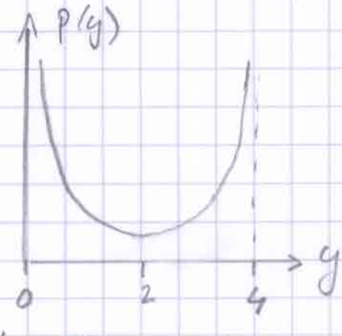
$$\Rightarrow dx = \frac{2}{\sqrt{y(4-y)}}\, dy \quad , \quad x=0 \leftrightarrow y=0\ ,\ x=\pi \leftrightarrow y=4$$

$$\frac{1}{n}\sum_{k=1}^{n} f(\lambda_k^{(n)}) \xrightarrow[n\to\infty]{} \int_0^4 f(y)\,\frac{2}{\pi\sqrt{y(4-y)}}\, dy$$

$p(y) = \dfrac{2}{\pi\sqrt{y(4-y)}}\cdot 1_{[0,4]}(y)$ limiting eigenvalue distribution



limiting eigenvalue profile



limiting eigenvalue distribution

<u>NB</u>: it is therefore possible to talk about "eigenvalue distribution" even for deterministic matrices!

Further generalizations of the theorem

1. If the sequence $(t_\ell, \ell \in \mathbb{Z})$ is not of finite order but satisfies still $\sum_{\ell \in \mathbb{Z}} |t_\ell| < \infty$, 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 $\sum_{\ell \in \mathbb{Z}} |t_\ell|^2 < \infty$ 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 $f(y) = 1_{y \le t}$ 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

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1_{\lambda_k^{(n)} \le t} = \frac{1}{2\pi} \int_0^{2\pi} 1_{g(x) \le t}\, dx \quad \forall t \in [a,b]$$

Let us define $F_n(t) := \frac{1}{n} \sum_{k=1}^{n} 1_{\lambda_k^{(n)} \leq t}$

and $F(t) := \frac{1}{2\pi} \int_0^{2\pi} 1_{g(x) \leq t} \, dx$. $\quad t \in [a,b]$

Note that

- $F_n(t) = \frac{1}{n} \# \{ k : \lambda_k^{(n)} \leq t \}$ is the proportion of eigenvalues of $T^{(n)}$ less than or equal to t
- by the same change of variable as above ($y = g(x)$)

$$F(t) = \int_a^t p(y) \, dy$$

Both F_n and F are therefore cumulative distribution functions$^{(*)}$, and the theorem says that

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

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

(*) 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 $[a,b]$)