Eigenvalues of Toeplitz Determinants

Graham Coppin

May 1990

Degree awarded with distinction 12 December 1990.

A research report submitted to the Faculty of Science, University of the Witwatersrand, in partial fulfilment of the degree of Master of Science.

Declaration

I declare that this research report is my own, unaided work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It [illegible] not been submitted before for any degree or examination in any other U[illegible]ty.

23rd day of [illegible], 1990

I am indebted to my supervisor Professor
Doron Lubinsky for his invaluable assistance
and unfailing confidence in my abilities.

This is dedicated to Wayne who would have liked to have seen a few more numbers.

Abstract

The Toeplitz form is a most useful and important tool in many areas of applied mathematics today including signal processing, time-series analysis and prediction theory. It is even used in quantum mechanics in Ising model correlation functions.

The eigenvalues of Toeplitz forms provide invaluable information and to this end it has become common to study their asymptotic nature. Much of the groundwork in this area is due to Grenander and Szegö whose seminal text the author will refer to extensively. The first part of this report is devoted to a restating and proof of their classical result on the asymptotic behaviour of the eigenvalues of a Toeplitz form. Following this a brief overview is given of important results obtained more recently which develop on the theme set by Grenander and Szegö.

The second part deals with the class of circulant matrices which is properly contained in the class of Toeplitz matrices and whose members are proving invaluable in the solving of the symmetric positive definite Toeplitz system $Ax = b$. We consider the gains made in this field by several authors towards more computationally efficient methods of solving the above system.

Contents

Chapter 1

Preliminaries

Otto Toeplitz was born in Breslau, Germany (now Wroclaw in Poland) and attended university there until moving to Göttingen, then the centre of world mathematics, where he joined the group around Hilbert and soon began working on problems involving infinite matrices and the associated systems of linear equations in infinitely many variables. Together with Ernst Hellinger he wrote some important papers in the years immediately prior to the Great War: in particular, they obtained one of the first examples of the uniform boundedness principle (Banach-Steinhaus theorem). At the same time Toeplitz introduced important matrices for summability and he contributed to the theory of almost periodic functions. Much of his motivation for studying Toeplitz matrices came from problems involving series and complex function theory.

Around 1913, Toeplitz moved to Kiel University to become the Dean of the Faculty of Philosophy where he became increasingly involved in the history of mathematics and its teaching. He wrote a calculus textbook and together with the number theorist Hans Rademacher, he wrote a very successful elementary problem book for a non-mathematical audience.

In the late 1920s he co-published (with Hellinger) a fundamental article on the theory of integral equations and then began a collaboration with G. Köthe that resulted in several basic contributions to the theory of sequence spaces.

In 1937 he became the first to identify the dual space of a space of analytic functions: in 1939 he moved to Jerusalem where he died in 1940.

His influence was such that by the 1920's the terms "Toeplitz form" and "Toeplitz matrix" were in common use. In this paper we shall examine only a fraction of the rich legacy left us by this man. It will be expedient to define certain concepts that play important roles in our investigation:

1. Suppose that $\{a_n\}_{n=1}^{\infty}$ is a sequence of real numbers satisfying
$$\lim_{n\to\infty} \quad a_n^{1/n} = r \tag{1.1}$$
where r is some real number. Then the sequence is said to satisfy an nth root asymptotic. A slightly stronger asymptotic is the ratio asymptotic
$$\lim_{n\to\infty} \quad \frac{a_n}{a_{n-1}} = r. \tag{1.2}$$
The strongest asymptotic is called a power or Szegö asymptotic and occurs when
$$\lim_{n\to\infty} \quad \frac{a_n}{r^n} = c \tag{1.3}$$
where c is real. It is important to note that
$$(1.3) \Rightarrow (1.2) \Rightarrow (1.1)$$
but that in general
$$(1.1) \not\Rightarrow (1.2) \not\Rightarrow (1.3).$$
In our investigation we shall encounter each of the three types.

2. Let $n \in N$ and suppose that $\{a_j^{(n)}\}_{j=1}^{n+1}$ and $\{b_j^{(n)}\}_{j=1}^{n+1}$ are sequences of real numbers satisfying
$$|\, a_j^{(n)} \,| < K, \quad |\, b_j^{(n)} \,| < K$$
where K is independent of j and n. Then the sequences $a_j^{(n)}$ and $b_j^{(n)}$, $n \to \infty$, are said to be equally distributed in the interval $[-K, K]$ if the following holds: Let $F(t)$ be an arbitrary continuous function in the interval $[-K, K]$; we have then
$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n+1}[F(a_j^{(n)}) - F(b_j^{(n)})]}{n+1} = 0. \tag{1.4}$$
This limit relation will hold for all continuous functions $F(t)$ if it holds for certain special sets of continuous functions, two of which are:

(a) (i) $F(t) = t^s$, $\quad s = 0, 1, 2, \ldots$. We have then

$$\lim_{n\to\infty} \frac{\sum_{j=1}^{n+1}[(a_j^{(n)})^s - (b_j^{(n)})^s]}{n+1} = 0. \tag{1.5}$$

(b) (ii) $F(t) = \log(1 + zt)$ where z is real and $|z| < K^{-1}$. In this case we have

$$\lim_{n\to\infty} \left\{ \frac{(1 + za_1^{(n)})(1 + za_2^{(n)})\ldots(1 + za_{n+1}^{(n)})}{(1 + zb_1^{(n)})(1 + zb_2^{(n)})\ldots(1 + zb_{n+1}^{(n)})} \right\}^{\frac{1}{n+1}} = 1. \tag{1.6}$$

3. Let $f(x)$ be a real-valued function of the class L, and let

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{1.7}$$

be its Fourier series, where

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-inx} f(x) dx; \quad c_{-n} = \overline{c_n}. \tag{1.8}$$

We recall that $T_n(u, v)$ is called a Hermitian form if

(a) $T_n(au_1 + bu_2, v) = af(u_1, v) + bf(u_2, v)$

(b) $T_n(u, v) = \overline{T_n(v, u)}$,

where a, b are arbitrary complex constants. Then it is easy to see that

$$T_n(u, u) = T_n(u) = \sum c_{k-j} u_j \overline{u_k} \quad j, k = 0, 1, \ldots, n \tag{1.9}$$

is Hermitian. We shall call it the nth Toeplitz form associated with the function $f(x)$.

We have

$$\begin{aligned} T_n &= \sum_{j,k=0}^{n} \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k-j)x} u_j \overline{u_k} f(x) dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j,k=0}^{n} e^{i(j-k)x} u_j \overline{u_k} f(x) dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=0}^{n} e^{ijx} u_j \sum_{k=0}^{n} \overline{e^{ikx} u_k} f(x) dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (u_0 + u_1 e^{ix} + \cdots + u_n e^{inx}) \overline{(u_0 + u_1 e^{ix} + \cdots + u_n e^{inx})} f(x) dx \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |u_0 + u_1 e^{ix} + \cdots + u_n e^{inx}|^2 f(x) dx. \end{aligned} \tag{1.10}$$

The $(n+1)$x$(n+1)$ matrix $(c_{j-k})_{j,k=0}^{n}$ is called the nth Toeplitz matrix associated with the function f and its determinant is denoted by $D_n(f)$.

4. Define $\Phi_0(z), \ldots, \Phi_n(z)$ to be the system of polynomials of the complex variable z which are orthonormal on the unit circle $|z| = 1$, $z = e^{ix}$ with the weight $\frac{1}{2\pi} f(x)dx$. The polynomials satisfy the following conditions:

   (a) $\Phi_n(z)$ is a polynomial of degree n whose leading coefficient defined to be k_n is real and positive;

   (b) $\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_n(z)\overline{\Phi_m(z)} f(x)dx = \delta_{nm}, \quad z = e^{ix}, \quad n, m = 0, 1, 2, \ldots$

   We mention that the system $\{\Phi_n(z)\}$ is uniquely determined by the above conditions.

   The polynomials may be represented explicitly:

$$\Phi_n(z) = (D_{n-1}D_n)^{-\frac{1}{2}} \det \begin{bmatrix} c_0 & c_{-1} & c_{-2} & \cdots & c_{-n} \\ c_1 & c_0 & c_{-1} & \cdots & c_{-n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & c_{n-3} & \cdots & c_{-1} \\ 1 & z & z^2 & \cdots & z^n \end{bmatrix}. \tag{1.11}$$

   This representation is easily derived by noting that the system of polynomials defined on the right-hand side of (1.11) satisfies the conditions so that by uniqueness they must be identical to $\{\Phi_n(z)\}$.
   From (1.11) we derive $k_n = (D_{n-1}/D_n)^{\frac{1}{2}}$.

Chapter 2

Classical results on eigenvalues of Toeplitz forms

2.1 Introduction

The eigenvalues of the Toeplitz form $T_n(f)$ are defined as the roots of the characteristic equation det $T_n(f-\lambda)=0$ and we denote them by

$$\lambda_1^{(n)}, \lambda_2^{(n)}, \cdots, \lambda_{n+1}^{(n)}.$$

Since each Toeplitz form is Hermitian it is clear that the eigenvalues are real. We shall investigate the asymptotic distribution of the eigenvalues as $n \to \infty$. We assume that $m \leq f(x) \leq M$ for some real m, M. Then (1.10) implies

$$m \leq T_n(f) \leq M \tag{2.1}$$

if we assume that $\{u_j\}_{j=0}^{n}$ are subjected to the condition

$$I = \sum_{j=0}^{n} |\, u_j \,|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\, u_0 + u_1 e^{ix} + \cdots + u_n e^{inx} \,|^2 \, dx = 1.$$

Thus we have

$$m \leq \lambda_k^{(n)} \leq M \qquad\qquad k = 1, 2, \ldots, n+1 \tag{2.2}$$

and if we order the eigenvalues in a nondecreasing way

$$\lambda_1^{(n)} \leq \lambda_2^{(n)} \leq \ldots \leq \lambda_{n+1}^{(n)}$$

then we find

$$\lambda_1^{(n)} \geq m, \qquad \lambda_{n+1}^{(n)} \leq M. \tag{2.3}$$

2.2 The Distribution Theorem

We are now ready to state Grenander and Szegös' distribution theorem:

Theorem 2.1 *Let $f(x)$ be a real-valued function of the class L. Denote by m and M the essential lower and upper bound of $f(x)$ respectively, where $m, M \in R$. If $F(\lambda)$ is any continuous function defined in $[m, M]$ then*

$$\lim_{n\to\infty} \frac{F(\lambda_1^{(n)}) + F(\lambda_2^{(n)}) + \cdots + F(\lambda_{n+1}^{(n)})}{n+1} = \frac{1}{2\pi}\int_{-\pi}^{\pi} F[f(x)]dx. \tag{2.4}$$

Here L is the space of all complex-valued functions $F(x)$ which are Lebesgue measurable and for which $\int_{-\pi}^{\pi} \mid F(x) \mid dx$ exists. Recall that the essential lower bound of $f(x)$ is that largest number m such that $f(x) \geq m$ holds except on a set of measure zero. The essential upper bound is defined similarly.

Using the terminology of equal distributions we may restate the limit in 2.1:

Proposition 2.1 *The limit relation (2.4) holds if and only if the sets*

$$\{\lambda_j^{(n)}\} \quad \text{and} \quad \{\mathrm{f}(-\pi + \frac{2\mathrm{j}\pi}{\mathrm{n}+2})\}, \quad \mathrm{n} \to \infty$$

are equally distributed.

Proof:
Write (2.4) as

$$\lim_{n\to\infty} \sum_{j=1}^{n+1} \frac{F(\lambda_j^{(n)})}{n+1} - \frac{1}{2\pi}\int_{-\pi}^{\pi} F(f[x])dx = 0. \tag{2.5}$$

Then setting $x_j = -\pi + \dfrac{2j\pi}{n+2}, \quad j = 1, 2, \ldots, n+1$ we have

$$\int_{-\pi}^{\pi} F[f(x)]dx = \lim_{n\to\infty} \frac{2\pi}{n+1} \sum_{j=1}^{n+1} F[f(x_j)]$$

since

$$\sum_{j=1}^{n+1} F[f(x_j)](x_{j+1} - x_j) = \frac{2\pi}{n+2} \sum_{j=1}^{n+1} F[f(x_j)]$$

is a Riemann sum for the integral $\dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} F[f(x)]dx$ and

$$\begin{aligned}
&\frac{2\pi}{n+1} \sum_{j=1}^{n+1} F[f(x_j)] - \frac{2\pi}{n+2} \sum_{j=1}^{n+2} F[f(x_j)] \\
&= \frac{2\pi}{(n+1)(n+2)} \sum_{j=1}^{n+1} F[f(x_j)] \\
&= O(\frac{1}{n^2})O(n) \\
&= O(\frac{1}{n}) \\
&\to 0 \quad \text{as} \quad n \to \infty.
\end{aligned}$$

Then (2.5) becomes

$$\lim_{n\to\infty} \sum_{j=1}^{n+1} \frac{F(\lambda_j^{(n)})}{n+1} - \lim_{n\to\infty} \frac{1}{n+1} \sum_{j=1}^{n+1} F[f(x_j)] = 0$$

$$\lim_{n\to\infty} \sum_{j=1}^{n+1} \frac{F(\lambda_j^{(n)}) - F[f(x_j)]}{n+1} = 0$$

i.e., $\{\lambda_j^{(n)}\}$ and $\{-\pi + \frac{2j\pi}{n+2}\}, \quad n \to \infty$ are equally distributed.□

Corresponding to the two special equal distribution cases we have

1. $F(\lambda) = \lambda^s, \quad s = 0, 1, 2, \ldots$ so that

$$\lim_{n\to\infty} \frac{(\lambda_1^{(n)})^s + (\lambda_2^{(n)})^s + \cdots + (\lambda_{n+1}^{(n)})^s}{n+1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x)]^s dx. \qquad (2.6)$$

2. Assume $m > 0$ so that $\lambda_j^{(n)} \geq m > 0$. Then $D_n(f) = \lambda_1^{(n)} \lambda_2^{(n)} \lambda_3^{(n)} \cdots \lambda_{n+1}^{(n)} > 0$ and we have

$$\lim_{n\to\infty} [D_n(f)]^{\frac{1}{n+1}} = \exp\{\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(x) dx\}. \qquad (2.7)$$

Recalling the discussion on equal distribution we see that we need only prove (2.7) to show that the two sets are equally distributed and thus that (2.4) holds.

Lemma 2.1 *The polynomial $k_n\Phi_n(z)(z)$ minimizes the integral*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mid g(z) \mid^2 f(x)dx, \quad z = e^{ix} \tag{2.8}$$

where $g(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_0$ is an arbitrary polynomial of degree n whose leading coefficient is unity. The minimum itself is $k_n^{-2} = \dfrac{D_n}{D_{n-1}}$.

Proof: Write $g(z) = v_0\Phi_0(z) + v_1\Phi_1(z) + \ldots + v_n\Phi_n(z)$ where $v_0, v_1, \ldots, v_n$ are complex variables and $v_nk_n = 1$. With $z = e^{ix}$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mid g(z) \mid^2 f(x)dx = \mid v_0 \mid^2 + \mid v_1 \mid^2 + \ldots + \mid v_n \mid^2 \geq \mid v_n \mid^2 = k_n^{-2}.\Box$$

Lemma 2.2 *The minimum of $\frac{1}{2\pi}\int_{-\pi}^{\pi} \mid g(z) \mid^2 f(x)dx, \quad z = e^{ix}$ where $g(z) = a_nz^n + a_{n-1}z^{n-1} + \ldots + a_1z + 1$ is an arbitrary polynomial whose constant term is 1 is identical to the minimum in the previous lemma.*

Proof:

$$\begin{aligned}
& \min_{a_n=1} \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \mid a_nz^n + \ldots + a_0 \mid^2 f(x)dx, \quad z = e^{ix} \\
= \; & \min_{a_n=1} \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \mid z^n \mid^2 \mid a_n + a_{n-1}z^{-1} + \ldots + a_0z^{-n} \mid^2 f(x)dx, \quad z = e^{ix} \\
= \; & \min_{a_n=1} \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} (a_n + a_{n-1}z^{-1} + \ldots + a_0z^{-n})(\overline{a_n} + \overline{a_{n-1}}z + \ldots + \overline{a_0}z^n)f(x)dx, \quad z = e^{i} \\
= \; & \min_{\overline{a_n}} \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \mid \overline{a_n} + \overline{a_{n-1}}z + \ldots + \overline{a_0}z^n \mid^2 f(x)dx, \quad z = e^{ix} \\
= \; & \min_{b_0=1} \tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \mid b_0 + b_1z + \ldots + b_nz^n \mid^2 f(x)dx, \quad z = e^{ix}.\Box
\end{aligned}$$

Theorem 2.2 *Let μ_n denote the minimum of*

$$T_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mid u_0 + u_1z + \ldots + u_nz^n \mid^2 f(x)dx, \quad z = e^{ix}; n = 0, 1, 2, \ldots$$

under the side condition $u_0 = 1$. Then

$$\lim_{n\to\infty} \mu_n = G(f) = exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \log f(x)dx\}.$$

In the case where $\log f(x)$ is not L-integrable, $G(f)$, the geometric mean of f, has to be replaced by zero.

Proof: See Grenander and Szegö [12].□

Corollary 2.1

$$\lim_{n\to\infty}[D_n(f)]^{\frac{1}{n+1}} = exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(x)dx\}.$$

Proof:

We have shown $\mu_n = \frac{D_n}{D_{n-1}}$. Thus by theorem (2.2) we have

$$\lim_{n\to\infty}\frac{D_n}{D_{n-1}} = exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(x)dx\}.$$

This is a ratio asymptotic which implies the weaker nth root asymptotic:

$$\lim_{n\to\infty}D_n(f)^{\frac{1}{n}} = exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(x)dx\};$$

this of course implies

$$\lim_{n\to\infty}[D_n]^{\frac{1}{n+1}} = exp\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(x)dx\}.$$

Thus (2.7) is proved, which implies that $\{\lambda_j^{(n)}\}$ and $\{-\pi + \frac{2j\pi}{n+2}\}$ are equally distributed, which in turn assures us of the validity of Theorem 2.1.

2.3 Refinement of the Distribution Theorem

Theorem 2.1 may be refined to incorporate a power asymptotic, although some additional assumptions must be made:

Condition 2.1 *Let $f(x)$ be a positive function of the class C and assume that $f'(x)$ exists and satisfies a Lipschitz condition:*

$$|f'(x_1) - f'(x_2)| < K|x_1 - x_2|^{\alpha}, \quad K > 0, \quad 0 < \alpha < 1.$$

We denote by $g(z) = g(f,z)$ the analytic function associated with $f(x)$ satisfying
$| g(z) |^2 = f(x), \quad z = e^{ix}$. Explicitly we have

$$\log g(z) = h(z) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \log f(x) \frac{1 + ze^{ix}}{1 - ze^{-ix}} dx = \sum_{n=0}^{\infty} h_n z^n. \qquad (2.9)$$

$g(z)$ is regular and different from zero in the open unit circle $| z | < 1$. Under condition(2.1) it turns out that it is continuous in the closed unit circle $| z | \leq 1$ and

$$\frac{1}{\pi} \int \int \left| \frac{g'(z)}{g(z)} \right|^2 d\sigma = \frac{1}{\pi} \int \int | h'(z) |^2 d\sigma = \sum_{n=0}^{\infty} n \, | h_n |^2 \qquad (2.10)$$

is finite, the latter series being convergent; here $d\sigma$ denotes the area element of the unit circle.

Lemma 2.3 *Let $f(x) = [\Phi(x)]^{-1}$ where $\Phi(x)$ is a positive trigonometric polynomial of precise degree p. We have*

$$\frac{D_n(f)}{[G(f)]^{n+1}} = \exp\left\{ \frac{1}{\pi} \int \int \left| \frac{g'(z)}{g(z)} \right|^2 d\sigma \right\} \qquad (2.11)$$

where the integration on the right is extended over the unit circle $| z | \leq 1$.

Proof:

Since $\Phi(x)$ is of precise degree p we can write it in the form $| \delta(z) |^2, z = e^{ix}$ where $\delta(z)$ is a polynomial of the precise degree p not vanishing in the unit circle $| z | \leq 1$ and we may assume $\delta(0) > 0$. We then have

$$g(f,z) = [\delta(z)]^{-1}$$

and for the geometric mean we obtain

$$\begin{aligned} G(f) &= \exp\{\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(x) dx\} \\ &= \exp\{-\frac{1}{\pi} \int_{-\pi}^{\pi} \Re \log \delta(z) dx\}, \qquad z = e^{ix} \\ &= \exp\{-2\Re \log \delta(0)\} \\ &= [\delta(0)]^{-2} \qquad (2.12) \end{aligned}$$

since $\Re \log \delta(z)$ is a harmonic function throughout the unit circle $|z| \leq 1$.

We assert that

$$\frac{D_n(f)}{[G(f)]^{n+1}} = \frac{D_{p-1}(f)}{[G(f)]^p}, \quad n \geq p \tag{2.13}$$

since by Theorem 2.2 and the fact that $\{\mu_n\}$ are non-increasing we have

$$G(f) \leq \mu_n = \frac{D_n(f)}{D_{n-1}(f)} \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| \frac{u_0 + u_1 z + \ldots + u_n z^n}{\delta(z)} \right|^2 dx, \quad z = e^{ix};$$

and $u_0 = 1, u_1, \ldots, u_n$ are arbitrary complex. Let $n \geq p$, and choose

$$u_0 + u_1 z + \ldots + u_n z^n = \frac{\delta(z)}{\delta(0)}.$$

Then by (2.12)

$$G(f) \leq \frac{D_n(f)}{D_{n-1}} \leq [\delta(0)]^{-2} = G(f)$$

so that

$$\frac{D_n(f)}{D_{n-1}(f)} = G(f), \quad n \geq p.$$

Hence

$$\begin{aligned} D_{p-1}(f)[G(f)]^{n-p+1} &= \frac{D_{p-1}(f)}{D_p(f)} \frac{D_p(f)}{D_{p+1}(f)} \cdots \frac{D_{n-1}(f)}{D_n(f)} D_n(f)[G(f)]^{n-p+1} \\ &= \frac{1}{[G(f)]^{n-p+1}} D_n(f)[G(f)]^{n-p+1} \\ &= D_n(f), \quad n \geq p \end{aligned}$$

and (2.13) follows.

We now need to evaluate the determinant $D_{p-1}(f)$ of the Toeplitz form:

$$\begin{aligned} T_{p-1}(f) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| \sum_{j=0}^{p-1} u_j e^{-ijx} \right|^2 f(x)dx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| \frac{u(z)}{\delta(z)} \right|^2 dx, \quad z = e^{ix} \end{aligned} \tag{2.14}$$

where

$$u(z) = \sum_{j=0}^{p-1} u_j z^{p-1-j}. \tag{2.15}$$

Denoting the distinct zeros of $\delta(z)$ by $z_0, z_1, \ldots, z_{p-1}$ we have

$$\delta(z) = \gamma(z_0 - z)(z_1 - z)\ldots(z_{p-1} - z) \tag{2.16}$$

where $|z_j| > 1$ and the product $\gamma z_0 z_1 \ldots z_{p-1}$ is positive.

By Lagrange's interpolation formula,

$$\frac{u(z)}{\delta(z)}=\sum_{j=0}^{p-1}\frac{u(z_j)}{\delta'(z_j)}\frac{1}{z-z_j}=\sum_{j=0}^{p-1}\frac{v_j}{z-z_j} \tag{2.17}$$

$$\text{where}\quad v_j=\frac{u(z_j)}{\delta'(z_j)}=\frac{1}{\delta'(z_j)}\sum_{k=0}^{p-1}u_k z_j^{p-1-k},\quad j=0,1,\ldots,p-1 \tag{2.18}$$

We now have a system representing a linear transformation of $\{u_j\}$ into $\{v_k\}$ with Jacobian

$$J=\frac{\partial(v_0,v_1,\ldots,v_{p-1})}{\partial(u_0,u_1,\ldots,u_{p-1})}=\frac{\prod_{\kappa<\lambda}(z_\kappa-z_\lambda)}{\delta'(z_0)\delta'(z_1)\ldots\delta'(z_{p-1})},\quad \kappa,\lambda=0,1,\ldots,p-1. \tag{2.19}$$

If one inserts for $\delta'(z_j)$ the expression obtained from (2.16) by differentiation and by the substitution $z=z_j$ we have another form of the Jacobian:

$$J=(-1)^{\frac{1}{2}p(p+1)}\gamma^{-p}\prod_{\kappa>\lambda}(z_\kappa-z_\lambda)^{-1}. \tag{2.20}$$

Grenander and Szegö now make use of (2.14) and (2.17) to transform $T_{p-1}(f)$:

$$\begin{aligned}
T_{p-1}(f) &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{u(z)}{\delta(z)}\right|^2dx,\quad z=e^{ix}\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\Big(\sum_{j=0}^{p-1}\frac{v_j}{z-z_j}\Big)\Big(\sum_{k=0}^{p-1}\frac{\overline{v_k}}{\overline{z-z_k}}\Big)dx,\quad z=e^{ix}\\
&= \sum_{j=0}^{p-1}\sum_{k=0}^{p-1}v_j\bar{v}_k\frac{1}{2\pi}\int_{|z|=1}\frac{dx}{(z-z_j)(\overline{z-z_k})},\quad z=e^{ix}.
\end{aligned}$$

Now

$$\begin{aligned}
&\frac{1}{2\pi}\int_{|z|=1}\frac{dx}{(z-z_j)(\overline{z}-\overline{z_k})}\\
&= \frac{1}{2\pi i}\int_{|z|=1}\frac{dz/z}{(z-z_j)(z^{-1}-\overline{z_k})}\\
&= \frac{1}{2\pi i}\int_{|z|=1}\frac{dz}{(z-z_j)(1-\overline{z_k}z)}\\
&= \lim_{z\to\overline{z_k}^{-1}}\frac{z-\overline{z_k}^{-1}}{(z-z_j)(1-\overline{z_k}z)}\\
&= \lim_{z\to\overline{z_k}^{-1}}\frac{\overline{z_k}z-1}{-(z-z_j)(\overline{z_k}z-1)}\frac{1}{\overline{z_k}}\\
&= \frac{1}{z_j-\overline{z_k}^{-1}}\frac{1}{\overline{z_k}}\\
&= \frac{1}{z_j\overline{z_k}-1}
\end{aligned}$$

where we have used residue theory and the fact that $|z_l|>1,\ l=0,1,\ldots,p-1$.

Now by Cauchy's determinant formula the determinant of $T_{p-1}(f)$ is

$$\begin{aligned}\det\left(\frac{1}{z_j\overline{z_k}-1}\right) &= \frac{1}{z_0 z_1\ldots z_{p-1}}\det\left(\frac{1}{\overline{z_k}-z_j^{-1}}\right)\\ &= \frac{\prod_{\kappa>\lambda}\mid z_\kappa - z_\lambda\mid^2}{\prod_{\kappa>\lambda}\mid \overline{z_\kappa} z_j - 1\mid}.\end{aligned}$$

Now from (2.20) $\mid J\mid^2 = \mid\gamma\mid^{-2p}\prod\mid z_\kappa - z_\lambda\mid^{-2}$, so we have

$$\begin{aligned}D_{p-1}(f) &= \mid J\mid^2 \det\left(\frac{1}{z_j\overline{z_k}-1}\right)\\ &= \frac{\mid\gamma\mid^{-2p}}{\prod_{\kappa>\lambda}\mid\overline{z_k}z_j - 1\mid}\end{aligned}$$

so that by (2.13),

$$\frac{D_n(f)}{[G(f)]^{n+1}} = \frac{\mid\gamma\mid^{-2p}}{\prod_{\kappa>\lambda}\mid\overline{z_k}z_j-1\mid}\frac{1}{[G(f)]^p}$$

$$= \frac{\mid z_0 z_1 \ldots z_{p-1}\mid^{2p}}{\prod_{\kappa>\lambda}\mid\overline{z_k}z_j - 1\mid},\quad j,k = 0,1,\ldots,p-1;\quad n\geq p, \tag{2.21}$$

by (2.12). Now

$$\begin{aligned} g'(z) &= \frac{-\delta'(z)}{[\delta(z)]^2}\\ \Rightarrow \frac{g'(z)}{g(z)} &= -\frac{\delta'(z)}{\delta(z)}\\ &= \frac{1}{z_0 - z} + \frac{1}{z_1 - z} + \cdots + \frac{1}{z_{p-1}-z}\end{aligned}$$

so that integrating over the unit circle $\mid z\mid\leq 1$, the right-hand expression in (2.11) becomes

$$\begin{aligned}&\frac{1}{\pi}\int\!\!\int \mid\frac{g'(z)}{g(z)}\mid^2 d\sigma\\ &= \sum_{j=0}^{p-1}\sum_{k=0}^{p-1}\frac{1}{\pi}\int\!\!\int\frac{d\sigma}{(z-z_j)(\overline{z}-\overline{z_k})}\end{aligned}$$

Now consider $\int\int \frac{d\sigma}{(z-z_j)(\overline{z}-\overline{z_k})}$. This can be written as

$$
\begin{aligned}
& \int_0^1 \int_{-\pi}^{\pi} \frac{r dr dx}{(re^{ix}-z_j)(re^{-ix}-\overline{z_k})} \\
= \; & \int_0^1 r dr \int_{|\zeta|=1} \frac{d\zeta/i\zeta}{(r\zeta - z_j)(r\zeta^{-1}-\overline{z_k})} \\
= \; & \frac{1}{i}\int_0^1 r dr \int_{|\zeta|=1} \frac{d\zeta}{(r\zeta - z_j)(r-\overline{z_k}\zeta)} \\
= \; & \frac{1}{i}\int_0^1 r dr \left[2\pi i \lim_{\zeta \to \frac{r}{\overline{z_k}}} \frac{\zeta - \frac{r}{\overline{z_k}}}{(r\zeta - z_j)(r-\overline{z_k}\zeta)} \right] \\
= \; & 2\pi \int_0^1 r dr \lim_{\zeta \to \frac{r}{\overline{z_k}}} \frac{-(r-\overline{z_k}\zeta)}{(r\zeta - z_j)(r-\overline{z_k}\zeta)\overline{z_k}} \\
= \; & 2\pi \int_0^1 r dr \left[\frac{-1}{(\frac{r^2}{\overline{z_k}} - z_j)\overline{z_k}} \right] \\
= \; & 2\pi \int_0^1 \frac{r}{z_j\overline{z_k} - r^2} dr \\
= \; & \pi \int_0^1 \frac{du}{z_j\overline{z_k} - u}.
\end{aligned}
$$

Now writing $z_j\overline{z_k} = s + it$ with s, t real,

$$\begin{aligned}
&\int_0^1 \frac{du}{z_j\overline{z_k} - u} + \int_0^1 \frac{du}{\overline{z_j}z_k - u} \\
&= 2\int_0^1 \frac{\Re(\overline{z_j}z_k - u)}{|z_j\overline{z_k} - u|^2} du \\
&= 2\int_0^1 \frac{s-u}{(s-u)^2+t^2} du \\
&= -\int_{s^2+t^2}^{(1-s)^2+t^2} \frac{dv}{v} \\
&= \log\left(\frac{s^2+t^2}{(1-s)^2+t^2}\right) \\
&= 2\log\left|\frac{\overline{z_j}z_k}{1-\overline{z_j}z_k}\right|.
\end{aligned}$$

Then

$$\begin{aligned}
&\exp\{\frac{1}{\pi}\int\int \left|\frac{g'(z)}{g(z)}\right|^2 d\sigma\} \\
&= \exp(\frac{1}{2}\sum_{j=0}^{p-1}\sum_{k=0}^{p-1}\left[\int_0^1 \frac{du}{z_j\overline{z_k} - u} + \int_0^1 \frac{du}{\overline{z_j}z_k - u}\right]) \\
&= \exp(\textstyle\sum_{j=0}^{p-1}\sum_{k=0}^{p-1}\log\left|\frac{\overline{z_j}z_k}{1-\overline{z_j}z_k}\right|) \\
&= \frac{|z_0 z_1 \ldots z_{p-1}|^{2p-1}}{\prod |z_j\overline{z_k} - 1|}.
\end{aligned}$$

Comparison with (2.21) establishes the lemma.□

We can now state the main result of this section:

Theorem 2.3 *(Refinement of Distribution Theorem)*
Let $f(x)$ satisfy Condition 2.1; denote by $G(f)$ the geometric mean of $f(x)$ and let $g(z) = g(f, z)$ be the analytic function associated with $f(x)$. Then

$$\lim_{n\to\infty} \frac{D_n(f)}{[G(f)]^{n+1}} = \exp\{\frac{1}{\pi}\int\int \left|\frac{g'(z)}{g(z)}\right|^2 d\sigma\}. \tag{2.22}$$

The integration on the right is extended over the unit circle.

Proof:

We approximate $f(x)$ by a function $f_1(x) = [\Phi(x)]^{-1}$ where $\Phi(x)$ is a positive trigonometric polynomial of precise degree p. Let $0 < \epsilon < 1$ and suppose

$$1 - \epsilon < \frac{f(x)}{f_1(x)} < 1 + \epsilon$$

Now

$$\begin{aligned}\frac{D_n(f)}{D_{n-1}(f)} &= \min_{g(z)=z^n+\dots} \frac{1}{2\pi}\int_{-\pi}^{\pi} \mid g(z) \mid^2 f(x)dx \\ &\leq (1+\epsilon) \min_{g(z)=z^n+\dots} \frac{1}{2\pi}\int_{-\pi}^{\pi} \mid g(z) \mid^2 f_1(x)dx \\ &= (1+\epsilon)\frac{D_n(f_1)}{D_{n-1}(f_1)}.\end{aligned}$$

Thus

$$(1-\epsilon)\frac{D_n(f_1)}{D_{n-1}(f_1)} \leq \frac{D_n(f)}{D_{n-1}(f)} \leq (1+\epsilon)\frac{D_n(f_1)}{D_{n-1}(f_1)}$$

where the left inequality is proved analogously.

Next we consider

$$\begin{aligned}\frac{D_n(f)}{D_m(f)} &= \prod_{j=m+1}^{n} \frac{D_j(f)}{D_{j-1}(f)} \\ &\leq \prod_{j=m+1}^{n} (1+\epsilon)\frac{D_j(f_1)}{D_{j-1}(f_1)} \\ &= (1+\epsilon)^{n-m}\frac{D_n(f_1)}{D_m(f_1)}\end{aligned} \tag{2.23}$$

$$\begin{aligned}\text{Now}\quad D_m(f) &= \det(c_{j-k}^{(f)})_{j,k=0}^{m} \\ \text{so}\quad D_0(f) &= \det(c_0^{(f)}) \\ &= c_0^{(f)} \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx \\ &\leq (1+\epsilon)\frac{1}{2\pi}\int_{-\pi}^{\pi} f_1(x)dx \\ &= (1+\epsilon)c_0^{(f_1)} \\ &= (1+\epsilon)D_0(f_1).\end{aligned}$$

Thus taking $m = 0$ in (2.23) we obtain

$$\begin{aligned}\frac{D_n(f)}{D_0(f)} &\leq (1+\epsilon)^n\frac{D_n(f_1)}{D_0(f_1)} \\ \Rightarrow D_n(f) &\leq (1+\epsilon)^n D_n(f_1)\frac{D_0(f)}{D_0(f_1)} \\ &\leq (1+\epsilon)^{n+1}D_n(f_1)\end{aligned}$$

Therefore

$$(1-\epsilon)^{n+1}D_n(f_1) \leq D_n(f) \leq (1+\epsilon)^{n+1}D_n(f_1) \tag{2.24}$$

where the left inequality is proved analogously.

Also we find that

$$(1-\epsilon)G(f_1) < G(f) < (1+\epsilon)G(f_1) \tag{2.25}$$

since

$$\begin{aligned} G(f) &= \exp[\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \log f(x)dx] \\ &\leq \exp[\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \log\{(1+\epsilon)f_1(x)\}dx] \\ &= (1+\epsilon)\exp[\tfrac{1}{2\pi}\textstyle\int_{-\pi}^{\pi} \log f_1(x)dx] \\ &= (1+\epsilon)G(f_1). \end{aligned}$$

The lower bound follows similarly.

Taking (2.24) and (2.25) into account we have

$$\left(\frac{1-\epsilon}{1+\epsilon}\right)^{n+1} \frac{D_n(f_1)}{[G(f_1)]^{n+1}} < \frac{D_n(f)}{[G(f)]^{n+1}} < \left(\frac{1+\epsilon}{1-\epsilon}\right)^{n+1} \frac{D_n(f_1)}{[G(f_1)]^{n+1}} \tag{2.26}$$

which will be of use to us later.

With the notation $z = re^{it}$, (2.9) becomes

$$h(re^{it}) = \frac{1}{4\pi}\int_{-\pi}^{\pi} \log f(x+t)\frac{1+re^{-ix}}{1-re^{-ix}}dx, \quad r<1, \tag{2.27}$$

so by differentiating with respect to t,

$$izh'(z) = \frac{1}{2}\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f'(x+t)}{f(x+t)}\frac{1+re^{-ix}}{1-re^{ix}}dx. \tag{2.28}$$

In view of standard theorems on boundary values of analytic functions, (see [12]) it turns out that $\Re[izh'(z)]$ has the limit $\dfrac{f'(t)}{f(t)}$ as $r \to 1$ and this satisfies a Lipschitz condition. By Privalov's theorem (see [12] for the reference) the same is true for $\Im[izh'(z)]$, for $h'(z)$ itself and thus also for $[g(z)]^{-1}$.

Now expand $[g(z)]^{-1}$ in a power series $\sum_{n=0}^{\infty} g_n z^n$, denoting its pth partial sum by $s_p(z)$. Grenander and Szegö refer us to general theorems on approximations which show us that $| s_p(e^{ix}) |^{-2} = f_p(x)$ approximates $f(x)$ with the degree of accuracy required, i.e.,

$$1 - \epsilon_p < \frac{f(x)}{f_p(x)} < 1 + \epsilon_p \quad \text{for} \quad \text{where} \quad \epsilon_p = O(p^{-1-\alpha} \log p), \ p \to \infty$$

so by (2.26),

$$\left(\frac{1-\epsilon_p}{1+\epsilon_p}\right)^p \frac{D_{p-1}(f_p)}{[G(f_p)]^p} \leq \frac{D_{p-1}(f)}{[G(f)]^p} \leq \left(\frac{1+\epsilon_p}{1-\epsilon_p}\right)^p \cdot \frac{D_{p-1}(f_p)}{[G(f_p)]^p}. \tag{2.29}$$

But $(1 \pm \epsilon_p)^p \to 1$ as $p \to \infty$ and by Lemma 2.3

$$\frac{D_{p-1}(f_p)}{[G(f_p)]^p} = \exp\{\frac{1}{2\pi} \int\int | \frac{s_p'(z)}{s_p(z)} |^2 \, d\sigma\}. \tag{2.30}$$

We need only show that the right hand expression in (2.30) tends to that in (2.22) as $p \to \infty$, i.e.,

$$\lim_{p\to\infty} \int\int | s_p(z) |^2 | \frac{s_p'(z)}{s_p(z)} |^2 \, d\sigma = \int\int \frac{1}{| g(z) |^2} | \frac{g'(z)}{g(z)} |^2 \, d\sigma.$$

This is clear since these integrals are equal to $\pi \sum_{n=1} n \mid g_n \mid^2$ and $\pi \sum_{n=1}^{\infty} n \mid g_n \mid^2$ respectively. Thus the proof is complete.□

2.4 Further results

We mention the following result due to Widom which is the culmination of the efforts of many authors and relaxes the conditions of the previous theorem. Let K be the space of functions $g : [-\pi, \pi] \to C$ having the Fourier expansion

$$g(\theta) \sim \sum_{k=-\infty}^{\infty} g_k e^{ik\theta}$$

with

$$||| g ||| \stackrel{def}{=} | g_0 | + \sum_{k=-\infty}^{\infty} | k || g_k |^2 < \infty.$$

Theorem 2.4 *Let $d\mu(\theta) \stackrel{def}{=} \mu'(\theta)d\theta$ be a non-negative measure on $[-\pi, \pi]$ such that $\mu' \in K$ and $\mu'^{-1} \in K$. Then*

$$\lim_{n\to\infty} \hat{D}_n(d\mu)/G[\mu']^n = \exp(\sum_{k=0}^{\infty} k(\log \mu')_k(log \mu')_{-k})$$

where $\{(\log \mu')_k\}_{-\infty}^{\infty}$ are the Fourier coefficients of $\log \mu'(\theta)$.

We mention that Widom's results are functional analytic in nature though retaining traces of Szegö's ideas.

Chapter 3

Eigenvalues of Toeplitz matrices associated with orthogonal polynomials

Let $d\alpha$ be a positive measure on the real line whose moments all exist. The support of $d\alpha$ is defined as

$$\text{supp}(d\alpha) = \{x \in R \quad | \quad \forall\epsilon > 0 : \alpha((x-\epsilon, x+\epsilon)) > 0\}. \tag{3.1}$$

and if $\text{supp}(d\alpha)$ is an infinite set then there exists a unique sequence of polynomials $p_n(x,\alpha) = k_n x^n + \dots \quad (n = 0, 1, 2, \dots)$ with $k_n > 0$ such that

$$\int_{-\infty}^{\infty} p_n(x,\alpha)p_m(x,\alpha)d\alpha(x) = \delta_{mn}.$$

These polynomials are said to be orthonormal with measure $d\alpha$. Let g be a real-valued α-measurable function such that for every $k \in N$,

$$\int_{-\infty}^{\infty} | x^k g(x) | \, d\alpha(x) < \infty;$$

then the infinite Toeplitz matrix $T(g,\alpha)$ associated with the orthogonal polynomials $p_n(x,\alpha)$ consists of the entries

$$[T(g,\alpha)]_{ij} = \int_{-\infty}^{\infty} p_i(x,\alpha)p_j(x,\alpha)g(x)d\alpha(x), \quad i,j = 0,1,2,\dots .$$

The truncated matrix $T_n(g,\alpha)$ contains the first n rows and columns of $T(g,\alpha)$:

$$T_n(g,\alpha) = [T(g,\alpha)]_{i,j=0}^{n-1}.$$

We introduce the modified truncated Toeplitz matrix

$$T_n^*(g,\alpha) = \left[\int_{-\infty}^{\infty} p_i(x,\alpha)p_j(x,\alpha)g(\frac{x}{c_n})d\alpha(x)\right]_{i,j=0}^{n-1}$$

where $\{c_n\}_{n=1,2,...}$ is a given sequence of positive numbers.

We denote the zeros of $p_n(x,\alpha)$ in increasing order by

$$x_{1,n} < x_{2,n} < \ldots < x_{n,n}$$

and the eigenvalues of $T_n^*(g,\alpha)$ which are real by

$$\Lambda_{1,n} \le \Lambda_{2,n} \le \ldots \le \Lambda_{n,n}.$$

Grenander and Szegö assert that if $\operatorname{supp}(d\alpha) = [-1,1]$ and $\log\alpha'(\cos\theta) \in L'$ and g is continuous in $[-1,1]$, then with $c_n = 1\forall n$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} G(\Lambda_{k,n}(g,d\alpha)) = \frac{1}{\pi}\int_{-1}^{1}\frac{G(g(x))}{(1-x^2)^{\frac{1}{2}}}dx$$

whenever G is continuous.

Nevai [17] has shown their proof to be incorrect since they used the Gauss-Jacobi mechanical quadrature with n nodes for polynomials of degree larger than $2n-1$. He relaxes the condition on g and reaches the same conclusion if g only belongs to L^∞ [18].

Theorem 3.1 *(Nevai)*
Let $supp(d\alpha) = [-1,1]$ and $\alpha'(x) > 0$ for almost every $x \in [-1,1]$ (the so-called Erdös-Turán condition). Assume that the Toeplitz matrix $T(g,d\alpha)$ is generated by an L^∞ function g. Let G be a continuous function in an interval containing the essential range of g. Then the eigenvalues Λ_{kn} of the truncated matrix $T_n(g,d\alpha)$ satisfy

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} G(\Lambda_{kn}) = \frac{1}{\pi}\int_{-1}^{1}\frac{G(g(x))}{(1-x^2)^{\frac{1}{2}}}dx.$$

Let us now introduce a sequence of discrete measures $\mu_n(n = 1,2,3,\ldots)$ on R by

$$\begin{cases} \mu_n(\{\frac{x_{j,n}}{c_n}\}) &= \frac{1}{n}; \\ \mu_n(A) &= 0 \text{ whenever } A \text{ contains no } \frac{x_{j,n}}{c_n}. \end{cases} \tag{3.2}$$

If there exists a sequence c_n such that the measures μ_n converge weakly to a probability mass μ then we say that the "contracted zeros" $\{\frac{x_{j,n}}{c_n} \mid j = 1, 2, \ldots, n\}$ are asymptotically distributed according to μ. We recall that weak convergence of a sequence of probability measures μ_n on R^k to a probability measure μ on R^k holds when for every bounded and continuous function f on R^k

$$\lim_{n \to \infty} \int_{R^k} f(x) d\mu_n(x) = \int_{R^k} f(x) d\mu(x).$$

Since we are interested in the asymptotic distribution of the eigenvalues of $T_n^*(g, \alpha)$ we define discrete measures $\nu_n (n = 1, 2, 3, \ldots)$ by

$$\begin{cases} \nu_n(\{\Lambda_{j,n}\}) & = \frac{k}{n} \quad \text{if } \Lambda_{j,n} \text{ has multiplicity } k; \\ \nu_n(A) & = 0 \quad \text{if } A \text{ contains no } \Lambda_{j,n}. \end{cases} \tag{3.3}$$

Van Assche's main theorem [21] connects the weak convergence of μ_n and of ν_n.

Theorem 3.2 *(van Assche)*
Suppose there exists a sequence $\{c_n\}$ such that the measures μ_n defined in (3.2) have a weak limit μ and such that $(\frac{1}{n})(k_{n-1}/k_n)$ is $o(\sqrt{n})$. If g is a bounded measurable function whose points of discontinuity form a set of μ-measure zero, then the measures ν_n as defined in (3.3) will converge weakly to μg^{-1} for which

$$\mu g^{-1}(A) = \mu(g^{-1}(A))$$

for every Borel set A.

Chapter 4

Circulant matrices

4.1 Introduction

A circulant is a special type of Toeplitz matrix exhibiting periodicity which connects it with Fourier analysis and group theory. Circulant matrices have applications in physics, image processing, probability and statistics, numerical analysis, number theory and geometry. We shall use them to give an alternative proof of Theorem 2.1 and in an investigation of recent developments in solving the algorithm $Ax = b$ where A is a Toeplitz matrix.

Suppose that C is an nxn matrix having the following structure:

$$\begin{aligned} C &= \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-3} & c_{n-2} \\ \vdots & & & \ddots & & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_0 \end{bmatrix} \\ &= [c_{(k-j) \bmod n}]_{j,k=0}^{n-1}. \end{aligned}$$

Then C is called a circulant matrix. The symbol or representer of C is

$$f(z) = \sum_{j=0}^{n-1} c_j z^j.$$

We note that

1. C is determined by its first row;
2. C is constant down its diagonals and is thus a Toeplitz matrix;
3. If C_1 and C_2 are circulant matrices and α, β are complex constants then $\alpha C_1 + \beta C_2$ is also circulant.

4.2 Eigenvalues and eigenvectors of circulant matrices

Let $\epsilon^n = 1$ and define x to be the vector

$$x = \begin{bmatrix} 1 \\ \epsilon \\ \epsilon^2 \\ \vdots \\ \epsilon^{n-1} \end{bmatrix}$$

We have the following result:

Theorem 4.1 *$Cx = f(\epsilon)x$, that is to say $f(\epsilon)$ is an eigenvalue of C with eigenvector x.*

Proof: The first component of Cx is

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon^2 \\ \vdots \\ \epsilon^{n-1} \end{bmatrix} = \sum_{j=0}^{n-1} c_j \epsilon^j = f(\epsilon) = f(\epsilon).1$$

Thus the first component of Cx equals the first component of $f(\epsilon)x$.

The second component of Cx is

$$\begin{aligned}
&\begin{bmatrix} c_{n-1} & c_0 & c_1 & \dots & c_{n-2} \end{bmatrix} \begin{bmatrix} 1 \\ \epsilon \\ \epsilon^2 \\ \vdots \\ \epsilon^{n-1} \end{bmatrix} \\
&\quad = c_{n-1} + c_0\epsilon + c_1\epsilon^2 + \dots + c_{n-2}\epsilon^{n-1} \\
&\quad = \epsilon(c_{n-1}\epsilon^{n-1} + c_0 + c_1\epsilon + \dots + c_{n-2}\epsilon^{n-2}) \\
&\quad = f(\epsilon).\epsilon
\end{aligned}$$

Thus the second components of Cx and $f(\epsilon)x$ are equal.

Continuing in this fashion we see that the jth component of Cx equals the jth component of $f(\epsilon)x$. □

We note that the eigenvector x is independent of $c_0, \dots, c_{n-1}$. Letting $\epsilon_j = e^{2\pi i(j-1)/n}, \quad j = 1, 2, \dots, n$ denote the distinct nth roots of unity, we set

$$x_j = \begin{bmatrix} 1 \\ \epsilon_j \\ \epsilon_j^2 \\ \vdots \\ \epsilon_j^{n-1} \end{bmatrix}$$

so that by Theorem 4.1

$$Cx_j = f(\epsilon_j)x_j. \tag{4.1}$$

Definition 4.1 *We define the* **Fourier matrix** *to be*

$$F = [x_1 \quad x_2 \quad \dots \quad x_n] = [\epsilon_k^{j-1}]_{j,k=1}^n.$$

Then (4.1) becomes

$$CF = FD, \quad D = \begin{bmatrix} f(\epsilon_1) & & \\ & \ddots & \\ & & f(\epsilon_n) \end{bmatrix}. \tag{4.2}$$

Lemma 4.1 *$\overline{F}^T F = nI$ so that F is non-singular and $n^{-\frac{1}{2}}F$ is orthogonal.*

Proof:

$$\begin{aligned}\overline{F}^T F &= [\overline{\epsilon_j}^{l-1}]_{j,l=1}^n [\epsilon_k^{l-1}]_{l,k=1}^n \\ &= [\sum_{l=1}^n (\overline{\epsilon_j}\epsilon_k)^{l-1}]_{j,k=1}^n.\end{aligned}$$

If $j = k$ then

$$\overline{\epsilon_j}\epsilon_k = | \epsilon_j |^2 = 1$$

and

$$\sum_{l=1}^n (\overline{\epsilon_j}\epsilon_k)^{l-1} = n.$$

If $j \neq k$ then

$$\overline{\epsilon_j}\epsilon_k = e^{2\pi i k - j/n} \neq 1$$

and

$$(\overline{\epsilon_j}\epsilon_k)^n = e^{2\pi i(k-j)} = 1,$$

so that we have

$$\sum_{l=1}^n (\overline{\epsilon_j}\epsilon_k)^{l-1} = \frac{1 - (\overline{\epsilon_j}\epsilon_k)^n}{1 - \overline{\epsilon_j}\epsilon_k} = \frac{1-1}{1 - \overline{\epsilon_j}\epsilon_k} = 0.\square$$

We can thus write

$$C = \frac{1}{n} F D F^{-1} = \frac{1}{n} F D \overline{F}^T \tag{4.3}$$

and we have proved the following result:

Theorem 4.2 *If C is a circulant with representer f then C admits the representation (4.3) with*

$$D = \begin{bmatrix} f(\epsilon_1) & & \\ & \ddots & \\ & & f(\epsilon_n) \end{bmatrix}$$

The eigenvalues of C are $f(\epsilon_j), j = 1, \ldots, n$ and

$$\det C = \prod_{j=1}^n f(\epsilon_j). \tag{4.4}$$

4.3 Further properties of circulants

Theorem 4.3 *(Characterization of circulants) Given any diagonal matrix*

$$D = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

where λ_i are arbitrary complex numbers, then (4.3) defines a circulant matrix.

Proof: Since the nth roots of unity $\epsilon_1, \ldots, \epsilon_n$ are distinct, there exists a unique Lagrange interpolation polynomial f of degree $\leq (n-1)$ such that

$$f(\epsilon_j) = \lambda_j, \quad j = 1, \ldots, n.$$

If $f(z) = \sum_{j=0}^{n-1} c_j z^j$ then taking this as the representer for a circulant matrix C we have by Theorem 4.2

$$\begin{aligned} C &= \frac{1}{n} F \begin{bmatrix} f(\epsilon_1) & & \\ & \ddots & \\ & & f(\epsilon_n) \end{bmatrix} F^{-1} \\ &= \frac{1}{n} F \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} F^{-1} \\ &= \frac{1}{n} F D F^{-1}. \square \end{aligned}$$

The above theorem shows that given any n complex numbers we can find a circulant with those numbers as eigenvalues. It is important to note that all circulants are diagonalised by the same matrix F.

Corollary 4.1 *If C_1, C_2 are circulants then so is C_1C_2. Moreover $C_1C_2 = C_2C_1$.*

Proof: We can write $C_1 = \frac{1}{n} F D_1 F^{-1}$ and $C_2 = \frac{1}{n} F D_2 F^{-1}$ for some diagonal matrices D_1, D_2. Then

$$C_1C_2 = \tfrac{1}{n} F D_1 F^{-1} \tfrac{1}{n} F D_2 F^{-1}$$

$$= \tfrac{1}{n}F(D_1D_2)F^{-1}$$

As D_1D_2 is diagonal, Theorem 4.3 shows that C_1C_2 is a circulant. Also

$$\begin{aligned} C_2C_1 &= \tfrac{1}{n}F(D_2D_1)F^{-1} \\ &= \tfrac{1}{n}F(D_1D_2)F^{-1} \\ &= C_1C_2.\square \end{aligned}$$

Corollary 4.2 *If C is a non-singular circulant then so is C^{-1}.*

Proof: If $C = \frac{1}{n}FDF^{-1}$ is invertible then D^{-1} exists and

$$C^{-1} = \frac{1}{n}FD^{-1}F.\square$$

We note the following:

1. The $n \times n$ vector space of circulants is a commutative ring (with respect to multiplication) whose units form an abelian group.
2. In multiplying or inverting circulants we need only compute the first row of the product or inverse.

4.4 Hadamard's inequality

Suppose $B = [b_{ij}]_{i,j=1}^{n}$ is an $n \times n$ matrix. Then

$$|\det B| \leq \prod_{j=1}^{n}[\sum_{k=1}^{n} |b_{jk}|^2]^{\frac{1}{2}}$$

Now if all the entries of the matrix have modulus which not exceeding unity we have

$$|\det B| \leq \prod_{j=1}^{n} n^{\frac{1}{2}} = n^{\frac{n}{2}}.$$

Theorem 4.4 *Let ϵ be a primitive nth root of unity where n is odd and let C be the circulant defined by*

$$c_j = \epsilon^{j(j+1)/2} \quad j = 0, 1, \ldots, n-1.$$

Then

$$| \det C | = n^{\frac{n}{2}}.$$

Proof: For a proof of this result see [16]. □

The above result shows that Hadamard's inequality is sharp for circulants.

5 The Distribution Theorem Revisited

We suppose that $f(x)$ is a real-valued measurable function bounded by M say. Then we may associate with f a certain $n \times n$ Toeplitz matrix $M_n(f)$:

$$M_n(f) = \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(k-j)x} f(x)dx\right]_{j,k=0}^{n-1}.$$

We introduce the Cesàro sums

$$f_p(x) = \sum_{j=-p}^{p}\left(1 - \frac{|j|}{p}\right) c_j e^{-ijx} = \sum_{j=-\infty}^{\infty} c'_j e^{-ijx},$$

where

$$c_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ijx} f(x)dx.$$

Now we know [13] that $| f_p(x) | \leq \sup_x f(x) \leq M < \infty$.

Let $\lambda_j = f_p(\frac{2\pi(j-1)}{n})$ $j = 1, \ldots, n$. Then by Theorem 4.3 we know that

$$C_n = \frac{1}{n} F \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} F^{-1}$$

is a circulant matrix with entries

$$
\begin{aligned}
[C_n]_{jk} &= n^{-1} \sum_{m=1}^{n} e^{2\pi i(j-1)(m-1)/n} \lambda_m e^{-2\pi i(m-1)(k-1)/n} \\
&= n^{-1} \sum_{m=1}^{n} e^{2\pi i(m-1)(j-k)/n} \lambda_m \\
&= n^{-1} \sum_{m=1}^{n} e^{2\pi i(m-1)(j-k)/n} f_p(\frac{2\pi(m-1)}{n}) \\
&= n^{-1} \sum_{m=1}^{n} e^{2\pi i(m-1)(j-k)/n} \sum_{l=-\infty}^{\infty} c_l' e^{-2\pi il(m-1)/n} \\
&= n^{-1} \sum_{l=-\infty}^{\infty} \sum_{m=1}^{n} c_l' e^{2\pi i(m-1)(j-k-l)/n}.
\end{aligned}
$$

Now $\sum_{m=1}^{n} z^{m-1} = 0$ if z is a primitive nth root of unity $\neq 1$, i.e.,

$$\sum_{m=1}^{n} e^{2\pi i(m-1)k/n} = n\delta_{k \ \mathrm{mod} n,0}.$$

Thus $\sum_{m=1}^{n} e^{2\pi i(m-1)(j-k-l)/n} = 0$ unless $j-k-l = \nu n, \nu = 0, \pm 1, \pm 2, \ldots.$

Hence $\sum_{m=1}^{n} e^{2\pi i(m-1)(j-k-l)/n} = 0$ unless $l = j-k+\nu n, \nu = 0, \pm 1, \pm 2, \ldots.$

Thus

$$
\begin{aligned}
[C_n]_{jk} &= n^{-1} \textstyle\sum_{\nu=-\infty}^{\infty} c_{j-k+\nu n}' \sum_{m=1}^{n} e^{2\pi i(m-1)(\nu n)/n} \\
&= \textstyle\sum_{\nu=-\infty}^{\infty} c_{j-k+\nu n}'
\end{aligned}
$$

If we write $K_n = [c_{k-j}']_{j,k=1}^{n}$ and define the matrix norm to be

$$
\begin{aligned}
| K_n | &= \left[\tfrac{1}{2\pi} \textstyle\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} | f_p(s-t) |^2 \, dsdt\right]^{\frac{1}{2}} \\
&= \left[\textstyle\int_{-\pi}^{\pi} | f_p(s) |^2 \, ds\right]^{\frac{1}{2}} \\
&= [2\pi \textstyle\sum_{\nu=-\infty}^{\infty} | c_\nu' |^2]^{\frac{1}{2}} \\
&= \left[2\pi \textstyle\sum_{\nu=-p}^{p} | c_\nu |^2 \, (1 - \frac{|\nu|}{p})^2\right]^{\frac{1}{2}}
\end{aligned}
$$

then it can be shown [12] that for large values of n

$$| C_n - K_n |^2 = 2n^{-1} \sum_{\nu=1}^{n} \nu \mid c_\nu' \mid^2 \leq 2pn^{-1} \sum_{\nu=1}^{\infty} \mid c_\nu \mid^2 = O(n^{-1})$$

and

$$| K_n - M_n(f) |^2 \leq 2n^{-1} \sum_{\nu=1}^{p} (\frac{\nu^2}{p^2}(p-\nu) | c_\nu |^2) + 2n^{-1} \sum_{\nu=p+1}^{n} (n-\nu) | c_\nu |^2 .$$

For a given $\epsilon > 0$ we can choose p so large that $\sum_{\nu=p+1}^{\infty} < \epsilon$. Then

$$| M_n(f) - K_n | \leq O(n^{-\frac{1}{2}}) + [O(n^{-1}) + 2\epsilon]^{\frac{1}{2}}$$

so that by choosing first p and then n sufficiently large the distance $| M_n(f) - C_n |$ can be made arbitrarily small.

Now C_n is Hermitian and bounded, $| L_n |$ and $\| L_n \| \leq M$ and we know by Theorem 4.3 that C_n has eigenvalues $f_p(\frac{2\pi(j-1)}{n}), j = 1, 2, \ldots, n$. In view of the continuity of $f_p(x)$ we have for any positive integer s

$$\lim_{n\to\infty} n^{-1} \sum_{j=1}^{n} \left[f_p(\frac{2\pi j}{n}) \right]^s = \frac{1}{2\pi} \int_{-\pi}^{\pi} [f_p(x)]^s dx$$

since the left-hand side is a Riemann sum for the right-hand integral. Also

$$\begin{aligned} | \int_{-\pi}^{\pi} ([f_p(x)]^s - [f(x)]^s) dx | \ &\leq sM^{s-1} \textstyle\int_{-\pi}^{\pi} | f_p(x) - f(x) | dx \\ &\leq (2\pi)^{\frac{1}{2}} sM^{s-1} \| f_p - f \| \\ &\to 0 \text{ as } p \to \infty. \end{aligned}$$

Thus given $\epsilon > 0$, we can choose first a p, and then an N_ϵ such that $n \geq N_\epsilon$ implies

$$| n^{-1} \sum_{j=1}^{n} (f_p(\frac{2\pi j}{n}))^s - \frac{1}{2\pi} \int_{-\pi}^{\pi} [f(x)]^s dx | < \epsilon.$$

Also if $\lambda_{n1}, \lambda_{n2}, \ldots, \lambda_{nn}$ are the eigenvalues of $M_n(f)$, then one can bound

$$n^{-1} [\sum_{j=1}^{n} \lambda_{nj}^s - \sum_{j=1}^{n} f_p(\frac{2\pi j}{n})^s]$$

in terms of $| M_n(f) - K_n |$ which, if we recall the first special case of equal distribution yields another proof of our (unrefined) distribution theorem.

4.6 Solving Toeplitz systems with the aid of circulants

In areas such as signal processing and time-series analysis one has to solve the system

$$A_n x = b$$

where A_n is an $n \times n$ Toeplitz matrix. Direct methods of solving based on Levinson's recursion formula [14] have been in use for many years and these require $O(n^2)$ complexity; faster algorithms have been developed [1],[10] : the stability of these algorithms for symmetric positive definite matrices is discussed in Bunch [4]. More recently Strang [19] proposed the preconditioned gradient method wherein the Toeplitz matrix is replaced by a circulant. The number of operations per iteration is of order $O(n \log n)$ as circulant systems are efficiently solved by the Fast Fourier Transform.

The ordinary iterative method is given by the recursive formula

$$C_n x_{k+1} = (C_n - A_n)x_k + b.$$

Thus with x_0 the initial point we let

$$C_n x_1 = (C_n - A_n)x_0 + b$$

and solve for x_1. Denoting x^* by the solution of the problem and $e_k = x^* - x_k$ by the error after k steps we have

$$\begin{aligned}
C_n e_k &= C_n x^* - C_n x_k \\
&= b + (C_n - A_n)x^* - b - (C_n - A_n)x_{k-1} \\
&= (C_n - A_n)e_{k-1} \\
\text{Thus } e_k &= C_n^{-1}(C_n - A_n)e_{k-1} \\
\text{Hence } e_k &= [C_n^{-1}(C_n - A_n)]^k e_0
\end{aligned}$$

A well-known result in numerical analysis asserts that for an arbitrary approximation x_0 the sequence $\{x_0, x_1, \ldots\}$ converges to a solution x^* iff

$$\| C_n^{-1}(C_n - A_n) \| < 1$$

or equivalently iff

$$\rho[C_n^{-1}(C_n - A_n)] < 1.$$

Here $\rho(B)$ denotes the spectral radius of the matrix B.

Thus the major issue in analyzing convergence is the eigenvalues of $C_n^{-1}(C_n - A_n)$ or those of $C_n^{-1}A_n$. It turns out that the extreme eigenvalues of $C_n^{-1}A_n$are not in close control yet the preconditioned gradient method can be effective since its convergence rate depends on all the eigenvalues λ_i of $C_n^{-1}A_n$and not exclusively on λ_1 and λ_n: convergence is fast, it appears, when the eigenvalues are clustered.

Each iteration of the preconditioned gradient method contains the periodic linear system with coefficient matrix C_n, the multiplication by A_n, and the two inner products which appropriately orthogonalize the directions d_j. Formally we may state the method as follows: starting from $x_0 = 0$ and $r_0 = b$,

$$\begin{aligned}
\text{Solve } C_n z_{j-1} &= r_{j-1},\\
\beta_j &= \frac{z_{j-1}^T r_{j-1}}{z_{j-2}^T r_{j-2}} \quad (\beta_1 = 0),\\
d_j &= z_{j-1} + \beta_j d_{j-1} \quad (d_1 = z_0),\\
\alpha_j &= \frac{z_{j-1}^T r_{j-1}}{d_j^T A d_j},\\
x_j &= x_{j-1} + \alpha_j d_j,\\
r_j &= r_{j-1} - \alpha_j A d_j.
\end{aligned}$$

The question now is how to choose a circulant "close" to A_n. The simplest construction (due to Strang [19]) is to copy the centre diagonals of A_n and bring them round to complete the circulant. Thus if

$$A_n = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_1 & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_{n-2} & \cdots & a_0 \end{bmatrix}$$

then with $n = 2m$ we define the circulant S_n by

$$S_n = \begin{bmatrix} a_0 & a_1 & \cdots & a_m & \cdots & a_1 \\ a_1 & a_0 & \cdots & a_{m-1} & \cdots & a_2 \\ \vdots & \vdots & & \ddots & & \vdots \\ a_m & & & & & \\ \vdots & & & & & \\ a_1 & & \cdots & & & a_0 \end{bmatrix}$$

More precisely the entries of S_n are given in the first row by

$$s_k = \begin{cases} a_k, & 0 \leq k \leq m \\ a_{n-k}, & m \leq k < n. \end{cases}$$

Chan and Strang [5] supposed that A_n is a finite section of a fixed singly infinite positive definite matrix A_∞ with associated function

$$f(\theta) = \sum_{k=-\infty}^{\infty} a_k e^{ik\theta}$$

which is real and positive and belongs to the Wiener class (the class of functions whose coefficients satisfy $\sum_{-\infty}^{\infty} a_k < \infty$) : this ensures that the sequence $\{a_k\}$ is in l^1. They then demonstrated that the circulants S_n and S_n^{-1} are uniformly bounded and positive definite for large n and that the eigenvalues of $S_n^{-1}A_n$ are clustered around 1.

Of course there are other ways to choose the circulant. Chan [8] proposed the following idea: With $B \equiv C_n - A_n$ we have

$$\begin{aligned} \rho(C_n^{-1}B) \leq \| C_n^{-1}B \| &= \| (I + A_n^{-1}B)^{-1}A_n^{-1}B \| \\ &\leq \| (I + A_n^{-1}B)^{-1} \| \, \| A_n^{-1}B \| \\ &\leq \frac{\| A_n^{-1}B \|}{1 - \| A_n^{-1}B \|} \end{aligned}$$

provided that $\| A_n^{-1}B \| < 1$ in some consistent matrix norm. Now the upper bound is a strictly increasing function of $\| A_n^{-1}B \|$. Chan proposed choosing C_n so that $\| A_n^{-1}B \|$ is minimised. Since $\| A_n^{-1}B \| \leq \| A_n^{-1} \| \, \| B \|$ one is naturally led to the following formulation for the construction of C_n:

$$\min_{C_n \text{ circulant}} \| C_n - A_n \|$$

Theorem 4.5 *Let*

$$A_n = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{-1} & a_0 & \cdots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{-(n-1)} & a_{-(n-2)} & \cdots & a_0 \end{bmatrix}$$

and

$$C_n = \begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ c_1 & \cdots & \cdots & c_0 \end{bmatrix}$$

and suppose that $\|\cdot\|_F$ *is the matrix norm defined by*

$$\| P - Q \|_F^2 = \sum_{i=0}^{n-1}\sum_{j=0}^{n-1}(p_{ij} - q_{ij})^2$$

where $P = [p_{ij}]$ *and* $Q = [q_{ij}]$ *are* $n \times n$ *matrices. Then the matrix* C_n *whose entries are given by*

$$c_k = \frac{ka_{-(n-k)} + (n-k)a_k}{n}, \quad k = 0, 1, \dots, n-1$$

minimises $\| C_n - A_n \|_F$. *The minimum is given by*

$$\begin{aligned}\min_{C_n \text{ circulant}} \| C_n - A_n \|_F^2 &= \sum_{i=0}^{n-1}(a_i - a_{-(n-i)})^2 \left(\frac{i(n-i)}{n}\right) \\ &\le \frac{n}{4}\sum_{i=0}^{n-1}(a_i - a_{-(n-i)})^2.\end{aligned}$$

Proof: We have

$$N = \| C_n - A_n \|_F^2 = \sum_{i=1}^{n-1}[i(c_i - a_{-(n-i)})^2 + (n-i)(c_i - a_i)^2].$$

Then

$$\frac{\partial N}{\partial c_i} = [2i(c_i - a_{-(n-i)}) + 2(n-i)(c_i - a_i)] \; i = 0, 1, \dots, n-1.$$

Setting $\frac{\partial N}{\partial c_i} = 0$ for each $i = 0, 1, \dots, n-1$ we have

$$2ic_i - 2ia_{-(n-i)} + 2(n-i)c_i - 2(n-i)a_i = 0$$

$$\Rightarrow c_i = \frac{ia_{-(n-i)} + (n-i)a_i}{n} \; i = 0, 1, \dots, n-1.$$

We then have

$$\begin{aligned}\| C_n - A_n \|_F^2 &= \textstyle\sum_{i=0}^{n-1}\left[i\left(\frac{ia_{-(n-i)}+(n-i)a_i}{n} - a_{-(n-i)}\right)^2 + (n-i)\left(\frac{ia_{-(n-i)}+(n-i)a_i}{n} - a_i\right)^2\right] \\ &= \textstyle\sum_{i=0}^{n-1}(a_i - a_{-(n-i)})^2\left[i\left(\frac{n-i}{n}\right)^2 + (n-i)\left(\frac{i}{n}\right)^2\right] \\ &= \textstyle\sum_{i=1}^{n-1}(a_i - a_{-(n-i)})^2\frac{i(n-i)}{n} \\ &\le \textstyle\frac{n}{4}\sum_{i=0}^{n-1}(a_i - a_{-(n-i)})^2. \square\end{aligned}$$

We note that C_n is symmetric iff A_n is symmetric and $C_n = A_n$ iff $a_i = a_{-(n-i)}$ $i = 1, 2, \dots, n-1$, i.e. iff A_n is itself circulant. We remark also that this

C_n is certainly different from the one proposed earlier by Strang. In numerical experiments for four different Toeplitz matrices the spectrum of $C_n A_n^{-1}$ was shown to lie within that of $S_n A_n^{-1}$ - hence the condition number of $C_n A_n^{-1}$ in the experiments was smaller.

Further research by Chan [6] resulted in his proving that if the generating function f is a positive function in the Wiener class then the spectra of C_n and S_n are asymptotically equal. In particular this means that for n sufficiently large C_n and C_n^{-1} are uniformly bounded in the l_2 norm and the eigenvalues of $C_n A_n^{-1}$ are clustered around 1:

Lemma 4.2 *Let the generating function f be a positive function in the Wiener class; then*

$$\lim_{n\to\infty} \rho(S_n - C_n) = 0.$$

Proof: Clearly $B_n \equiv S_n - C_n$ is a circulant matrix and it is easy to verify that its entries are given by

$$b_k = \begin{cases} \frac{k}{n}(a_k - a_{n-k}), & 0 \leq k \leq m, \\ \frac{n-k}{n}(a_{n-k} - a_k), & m \leq k < n. \end{cases}$$

Now by Theorem 4.1 the jth eigenvalue $\lambda_j(B_n)$ of B_n is given by

$$\begin{aligned} \lambda_j(B_n) &= \textstyle\sum_{k=0}^{n-1} b_k e^{\frac{2\pi ijk}{n}} \\ &= 2\textstyle\sum_{k=1}^{m-1} \frac{k}{n}(a_k - a_{n-k}) \cos\frac{2\pi jk}{n}. \end{aligned}$$

Then the spectral radius is bounded above as follows:

$$\rho(B_n) \leq 2\sum_{k=1}^{m-1} \frac{k}{n} \mid a_k \mid + 2\sum_{k=m+1}^{n-1} \mid a_k \mid .$$

f is a member of the Wiener class, thus given any $\epsilon > 0$ we can find an $M_1, M_2 > 0$ with $M_2 > M_1$, such that

$$\sum_{k=M_1+1}^{\infty} \mid a_k \mid < \frac{\epsilon}{6} \text{ and } \frac{1}{M_2}\sum_{k=1}^{M_1} k \mid a_k \mid < \frac{\epsilon}{6}.$$

Then with $m > M_2$ we are ensured that

$$\rho(B_n) < \frac{2}{M_2}\sum_{k=1}^{M_1} k \mid a_k \mid + 2\sum_{k=M_1+1}^{m-1} \mid a_k \mid + 2\sum_{k=m+1}^{\infty} \mid a_k \mid < \epsilon \; \Box.$$

It is known that S_n and S_n^{-1} are uniformly bounded. If f is bounded then the spectrum of S_n lies in $[f_{\min}, f_{\max}]$. Lemma 4.2 then gives us the following theorem:

Theorem 4.6 *Let f be a positive function in the Wiener class, then for all n sufficiently large, the circulant matrices C_n and C_n^{-1} are uniformly bounded in the l_2 norm. Moreover $\sigma(C_n)$ lies in $[f_{\min}, f_{\max}]$.*

Using the result which Chan and Strang had established earlier which asserted that the spectrum of $A_n - S_n$ is clustered around zero we are then able to state the following result:

Theorem 4.7 *Let f be a positive function in the Wiener class; then for all $\epsilon > 0$, there exist $M, N > 0$ such that, for all $n > N$, at most M eigenvalues of $A_n - S_n$ have absolute value larger than ϵ.*

Thus the spectrum of $C_n A_n^{-1}$ is clustered around 1 for sufficiently large n.

References

[1] G. Ammar and W. Gragg, Superfast solution of real positive definite Toeplitz Systems, *SIAM J. Matrix Anal. Appl.*, 9 (1988), 61-76.

[2] R. Bitmead and B. Anderson, Asymptotically fast solution of Toeplitz and related systems of equations, *Linear Algebra Appl.*, 34 (1980), 103-[illegible]6.

[3] R. Brent, F. Gustavson, and D. Yun, Fast solution of Toeplitz systems of equations and computations of Padé approximations, *J. Algorithms, 1 (1980)*, 259-295.

[4] J. Bunch, Stability of methods for solving Toeplitz systems of equations, *SIAM J. Sci. Statist. Comput.*, 6 (1985), 349-364.

[5] R. Chan and G. Strang, Toeplitz equations by conjugate gradients with circulant preconditioner, *SIAM J. Sci. Statist. Comput.*, 10 (1989), pp 104-119.

[6] R. Chan, The spectrum of a family of circulant preconditioned Toeplitz systems, *SIAM J. Numer. Anal.*, 26 (1989), 503-506.

[7] R. Chan, Circulant preconditioners for Hermitian Toeplitz systems, *SIAM J. Matrix Anal. Appl.*, 10 (1989), 542-550.

[8] T. Chan, An optimal circulant preconditioner for Toeplitz systems, *SIAM J. Sci. Statist. Comput.*, 9 (1988), 766-771.

[9] P.J. Davis, *Circulant Matrices*, John Wiley and Sons, Inc., New York 1979.

[10] F. de Hoog, A new algebra for solving Toeplitz Systems of equations, *Linear Algebra Appl.*, 88/89 (1987), 122-138.

[11] M.E. Fisher and R.E. Hartwig, Toeplitz determinants: Some Applications, Theorems, and Conjectures, *Adv. in Chem. Phys.*, 15 (1968), 333-353.

[12] U.Grenander and G.Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, CA, 1958.

[13] Y. Katznelson, *An Introduction to Harmonic Analysis*, Dover, New York, 1968.

[14] N. Levinson, The Wiener RMS (Root-Mean Square) Error Criterion in Filter Design and Prediction, *J. Math. Phys.*, 25 (1947), 261-278.

[15] D.S. Lubinsky, A Survey of General Orthogonal Polynomials for Weights on Finite and Infinite Intervals, *Acta Applicandae Mathematica*, 10 (1987), 237-296.

[16] D.S. Lubinsky and B. Shekhtman, Unimodular Toeplitz and Circulant Matrices with Large Determinant, preprint.

[17] P.G. Nevai, Orthogonal Polynomials, *Mem. Amer. Math. Soc, No 213, Amer. Math. Soc.*, Providence, R.I., 1979.

[18] P.G. Nevai, Eigenvalue Distribution of Toeplitz matrices, *Proc. Am. Math. Soc.*, 80 (1980), 247-253.

[19] G.Strang, A proposal for Toeplitz Matrix Calculation, *Stud. Appl. Math.*, 74 (1986), 171-176.

[20] G.Szegö, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publns., Vol. 23, Providence, 1939, 4th edition (reprinted), 1985.

[21] W. van Assche, Eigenvalues of Toeplitz Matrices Associated with Orthogonal Polynomials, *J. of Approx. Theory*, 51 (1987), 360-371.

[22] A. Zygmund, *Trigonometric Series* Second Edition Volumes I and II, Cambridge University Press, Cambridge, 1959.