A STUDY OF CICULANT AND TOEPLITZ MATRICES

A Project Report Submitted

in Partial Fulfilment of the Requirements

for the Degree of

MASTER OF SCIENCE

in

Mathematics and Computing

by

Akarshit kumar

(Roll No. 222123004)

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781039, INDIA

April 2024

CERTIFICATE

This is to certify that the work contained in this report entitled **"A STUDY OF CICULANT AND TOEPLITZ MATRICES"** submitted by **Akarshit kumar** (**Roll No: 222123004**) to Department of Mathematics, Indian Institute of Technology Guwahati towards the requirement of the course **MA699 Project** has been carried out by him under my supervision. It is also certified that the content of this report is based on the references in the Bibliography.

Guwahati - 781 039	(Dr. Rafikul Alam)
April 2024	Project Supervisor

ABSTRACT

The aim of this project is to study specialized matrices known as circulant matrices and Toeplitz matrices and investigate their properties. Circulant matrices possess a unique characteristic where each row is a cyclic shift of the previous row. We explore how circulant matrices are constructed and study their properties. We also delve into Toeplitz matrices, which have constant diagonals. Specifically, we look at Wiener class of Toeplitz matrices, which is used for study of Toeplitz matrices. We also discusses the behavior of these matrices as they grow larger (asymptotic behavior) which is used to study the behaviour of eigenvalues of matrices.

Contents

Chapter 1

Circulant Matrices

1.1 Introductory Properties

A Circulant Matrix or a Circulant of order n, is a square matrix of the form

$$C = \operatorname{circ}(c_1, c_2, \ldots, c_n) = \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ c_n & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_n & c_1 & \cdots & c_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_2 & c_3 & c_4 & \cdots & c_1 \end{bmatrix} \tag{1.1}$$

The elements of each row of C are identical to those of the previous row, but are moved one position to the right and wrapped around. The whole circulant is evidently determined by the first row or column. We may also write a circulant in the form

$$C = (c_{jk}) = (c_{k-j+1}), \quad \text{subscripts mod } n. \tag{1.2}$$

Theorem 1.1.1. *Let C be an $n \times n$ matrix. Then C is circulant if and only if $C\pi = \pi C$, where π is the circulant matrix defined as $\pi = circ(0, 1, 0, \ldots, 0)$.*

Definition 1.1.2. In view of the structure of the permutation matrices π^k, $k = 0, 1, \ldots, n-1$, it is clear that

$$\text{circ}(c_1, c_2, \ldots, c_n) = c_1 I + c_2 \pi + c_3 \pi^2 + \ldots + c_n \pi^{n-1}. \tag{1.3}$$

Thus, from Theorem 1.1.1, C is circulant if and only if $C = p(\pi)$ for some polynomial $p(z) = c_1 + c_2 z + \ldots c_n z^{n-1}$ associated with the n-tuple $\gamma = (c_1, c_2, \ldots, c_n)$. The polynomial $p_\gamma(z)$ will be called the representer of the circulant matrix. Thus,

$$C = \text{circ}(\gamma) = p_\gamma(\pi). \tag{1.4}$$

1.2 Diagonalization of Circulants

Definition 1.2.1. Let n be a fixed integer ≥ 1. Let $\omega = e^{2\pi i/n} = \cos\left(\frac{2\pi}{n}\right) + i \sin\left(\frac{2\pi}{n}\right)$. Define

$$\Omega = \text{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}). \tag{1.5}$$

Definition 1.2.2. The Fourier matrix of order n is given by

$$F^* = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \cdots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{n-1} & \omega^{2(n-1)} & \cdots & \omega^{(n-1)(n-1)} \end{bmatrix} \tag{1.6}$$

The Fourier matrix is unitary, it's eigenvalues are ± 1 and $\pm i$ and it has various important properties.

Theorem 1.2.3. *We have, $\pi = F^*\Omega F$, where F is the Fourier matrix.*

Proof. The j-th row of F^* is $\frac{1}{\sqrt{n}}(\omega^{(j-1)0}, \omega^{(j-1)1}, \ldots, \omega^{(j-1)(n-1)})$. Hence, the j-th row of $F^*\Omega$ is $\left(\frac{1}{\sqrt{n}}\omega^{(j-1)r} \cdot \omega^r\right)$ for $r = 0, 1, \ldots, n-1$ where $\omega = e^{2\pi i/n}$.

Now, the k-th column of F is $\left(\frac{1}{\sqrt{n}}\overline{z}^{(k-1)r}\right)$ for $r = 0, 1, \ldots, n-1$.

Thus, the (j,k)-th element of $\hat{F}\Omega F$ is

$$\sum_{r=0}^{n-1} \omega^{jr}\overline{\omega}^{(k-1)r} = \frac{1}{n}\sum_{r=0}^{n-1} \omega^{r(j-k+1)} \mod n = \begin{cases} 1 & \text{if } j \equiv k-1 \pmod{n} \\ 0 & \text{if } j \not\equiv k-1 \pmod{n} \end{cases}.$$

$\square$

Theorem 1.2.4. *If C is a circulant matrix, it is diagonalized by F. More precisely,*

$$C = F^*\Lambda F \tag{1.7}$$

where $\Lambda = \Lambda_C = diag(p_\gamma(1), p_\gamma(\omega), \ldots, p_\gamma(\omega^{n-1}))$.

Proof. From Equation (1.4), we have

$$C = \mathrm{circ}(\gamma) = p_\gamma(\pi).$$

Now from Theorem 1.2.3, $\pi = F^*\Omega F$, that is, we have

$$\begin{aligned}
C &= p_\gamma(F^*\Omega F) \\
&= c_1 + c_2(F^*\Omega F) + c_3(F^*\Omega F)^2 + \ldots + c_n(F^*\Omega F)^{n-1} \\
&= c_1 + c_2(F^*\Omega F) + c_3(F^*\Omega^2 F) + \ldots + c_n(F^*\Omega^{n-1} F) \\
&= F^*(c_1 + c_2\Omega + c_3\Omega^2 + \ldots + c_n\Omega^{n-1})F \\
&= F^* p_\gamma(\Omega) F \\
&= F^* \mathrm{diag}(p_\gamma(1), p_\gamma(\omega), \ldots, p_\gamma(\omega))F.
\end{aligned}$$

□

Now, the eigenvalues of circulant matrix C are

$$\lambda_j = p_\gamma(\omega^{j-1}), \quad j = 1, 2, \ldots, n. \tag{1.8}$$

and the columns of F^* are a universal set of eigenvectors for all circulant matrices.

Theorem 1.2.5. *Let* $\Lambda = diag(\lambda_1, \lambda_2, \ldots \lambda_n)$; *then* $C = F^*\Lambda F$ *is a circulant matrix.*

Proof. Consider the points $((1, \lambda_1), (\omega, \lambda_2), \ldots, (\omega^{n-1}, \lambda_n))$ By polynomial interpolation we can find a unique polynomial $r(z)$ of degree $\leq (n-1)$, $r(z) = d_1 + d_2 z + \ldots + d_n z^{n-1}$ such that $r(\omega^{j-1}) = \lambda_j, j = 1, 2, \ldots, n$.

Now, form $D = \mathrm{circ}(d_1, d_2, \ldots, d_n)$. Then $D = F^*\Lambda F = C$. Hence, C is circulant.

□

Definition 1.2.6. A matrix is called nondefective or semi-simple if and only if it is diagonalizable.

From (1.7) since all circulant matrices are diagonalizable, all circulants matrices are semi-simple.

1.3 Determinants of Circulant Matrices

The determinant of a square matrix is the product of its eigenvalues. Therefore from (1.8)

$$\det(C) = \det \operatorname{circ}(c_0, c_1, \ldots, c_{n-1}) = \prod_{j=1}^{n-1} p_\gamma(\omega^{j-1})$$

Definition 1.3.1. If $f(z) = a_0 z^m + a_1 z^{m-1} + \cdots + a_m, a_0 \neq 0$, and $g(z) = b_0 z^n + b_1 z^{n-1} + \cdots + b_n, b_0 \neq 0$ and have roots $\alpha_0, \alpha_1, \ldots, \alpha_{m-1}; \beta_0, \beta_1, \ldots, \beta_{n-1}$ respectively, the resultant $R(f, g)$ of f and g is defined as

$$R(f,g) = a_0^n g(\alpha_0) g(\alpha_1) \cdots g(\alpha_{m-1})$$

$$= a_0^n b_0^m \prod_{i=1}^{m-1} \prod_{j=1}^{n-1} (\alpha_i - \beta_j)$$

$$= (-1)^{mn} b_0^n f(\beta_0) f(\beta_1) \cdots f(\beta_{n-1})$$

$$= (-1)^{mn} R(g, f)$$

Thus, with $f(z) = z^n - 1$ and $g(z) = p_\gamma(z)$, we have

$$\det \operatorname{circ}(c_0, c_1, \ldots, c_{n-1}) = R(z^n - 1, p_\gamma(z))$$

$$= (-1)^{n(n-1)} R(p_\gamma; z^n - 1) = c_n^n \prod_{j=1}^{n-1} (\mu_j^n - 1)$$

where $\mu_0, \mu_1, \ldots, \mu_{n-1}$ are the roots of $p_\gamma(z)$.

In this way, the determinant of C can be expressed as the resultant of the two polynomials $z^n - 1$ and $p_\gamma(z)$.

1.4 Inversion of Circulant Matrices

Definition 1.4.1. The Moore-Penrose Inverse of a scalar λ,is given by

$$\lambda^+ = \begin{cases} \frac{1}{\lambda} & \text{if } \lambda \neq 0, \\ 0 & \text{if } \lambda = 0. \end{cases}$$

For $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots \lambda_n)$, set

$$\Lambda^+ = \text{diag}(\lambda_1^+, \lambda_2^+, \ldots, \lambda_n^+).$$

Theorem 1.4.2. *If C is the circulant $C = F^* \Lambda F$, then its Moore-Penrose generalized inverse is the circulant*

$$C^+ = F^* \Lambda^+ F.$$

Definition 1.4.3. (Another way of looking at Inverse of Circulant Matrices)

Let $C = \text{circ}(c_1, c_2, \ldots, c_n)$ and let $p(z) = c_1 + c_2 z + \ldots + c_n z^{n-1}$ be its representer. From (1.3)

$$C = p(\pi). \tag{1.9}$$

The last few coefficients in polynomial representer $p(z)$ of C may be zero. Assuming that $C \neq 0$, let us rewrite $p(z)$ as $p(z) = c_1 + c_2 z + \ldots + c_r z^{r-1}$ with $1 \leq r \leq n-1$ and $a_r \neq 0$.

Suppose that $\mu_1, \mu_2, \ldots, \mu_{r-1}$ are the zeros of the representer $p(z)$. Then $p(z) = a_r(z-\mu_1)(z-\mu_2)\ldots(z-\mu_{r-1})$, Hence

$$C = p(\pi) = a_r(\pi - \mu_1 I)(\pi - \mu_2 I)\ldots(\pi - \mu_{r-1} I). \tag{1.10}$$

This gives us a factorization of any circulant matrix into a product of circulants $(\pi - \mu_k I)$.

Now, suppose that C is non-singular. This is true if and only if $\lambda_j = p(\omega^{j-1}) \neq 0$, $j = 1, 2, \ldots, n$. This will be true if and only if $\mu_k^n \neq 1$ for $k = 1, 2, \ldots, r-1$, Hence from (1.10) we have,

$C^{-1} = a_r^{-1}(\pi - \mu_0 I)^{-1}(\pi - \mu_1 I)^{-1} \cdots (\pi - \mu_{r-1} I)^{-1}$.

Theorem 1.4.4. *Let $\mu^n \neq 1$. Then*

$$(\pi - \mu I)^{-1} = \frac{1}{1-\mu^n}[\mu^{n-1} I + \mu^{n-2}\pi + \mu^{n-3}\pi^2 + \ldots \pi^{n-1}]. \tag{1.11}$$

1.5 Additional Properties of Circulant Matrices

Definition 1.5.1 (Multiplication of Circulant Matrices)**.** Let $C_k, k = 1, 2, \ldots p$ be circulant matrices with $C_k = F^* \Lambda_k F$, Λ_k is diagonal. Then

$$\prod_{k=1}^{p} C_k = \prod_{k=1}^{p} (F^* \Lambda_k F) = F^* (\prod_{k=1}^{p} \Lambda_k) F. \tag{1.12}$$

From (1.12) it follows that eigenvalues of the product $C_1C_2\ldots C_p$ are the product of the eigenvalues of C_1, C_2, $\ldots$, C_p.

A special case of (1.12) is

$$C^k = F^*\Lambda^k F. \tag{1.13}$$

Definition 1.5.2. The rank of a diagonalizable matrix is equal to the number of its nonzero eigenvalues. Hence, if $C = F^*\Lambda F$, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, then $\text{RANK}(C) =$ number of the $\lambda's$ that are not zero.

Definition 1.5.3. Let $C = \text{circ}(c_1, c_2, \ldots c_n) = F^*\Lambda F$, $\Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then Trace of C is $\text{tr}(C) = \sum_{k=1}^{n} \lambda_k$.

Definition 1.5.4. Let C be a matrix whose characteristic polynomial is $P(\lambda) = (\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2}\ldots(\lambda - \lambda_s)^{\alpha_s}$ where $\lambda_1, \lambda_2, \ldots, \lambda_s$ are distinct and the integers $\alpha_k \geq 1$. Then the minimal polynomial of C has the form $m(\lambda) = (\lambda - \lambda_1)^{\beta_1}(\lambda - \lambda_2)^{\beta_2}\ldots(\lambda - \lambda_s)^{\beta_s}$ with $1 \leq \beta_j \leq \alpha_j$, $j = 1, 2, \ldots s$. Now A matrix is diagonalizable if its minimal polynomial has only simple zeros. Therefore, if C is a circulant matrix, then

$m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\ldots(\lambda - \lambda_s)$.

In other words, $m(\lambda)$ is a monic polynomial of minimal degree which has as its zeros all the distinct eigenvalues of C.

Chapter 2

The Asymptotic Behavior of Matrices

In this chapter, we review some definitions and theorems and proceed to a discussion of the asymptotic eigenvalues, product, and inverse behavior of sequences of matrices. The major use of these theorems is to relate the asymptotic behavior of a sequence of complicated matrices to that of a simpler, asymptotically equivalent sequence of matrices.

2.1 Convergence of Sequence of Matrices

Definition 2.1.1. Let $\{A_k\}$ be a sequence of $m \times n$ matrices and A be an $m \times n$ matrix. Then we say that $A_k \to A$ as $k \to \infty$ if $a_{ij}^{(k)} \to a_{ij}$ as $k \to \infty$ for all $i = 1, 2, \ldots, m$ and $j = 1, 2, \ldots, n$.

Definition 2.1.2. We say that $\sum_{k=1}^{\infty} A_k = A$ if $\lim_{k\to\infty} \sum_{j=1}^{k} A_j = A$.

Further, we say that $\prod_{k=1}^{\infty} A_k = A$ if $\lim_{k\to\infty} \prod_{j=1}^{k} A_j = A$.

2.2 Eigenvalues

By Schur's theorem, any matrix $A \in \mathbb{C}^n$ can be written as

$$A = URU^*, \tag{2.1}$$

where U is a unitary matrix (i.e., $U^* = U^{-1}$) and $R = (r_{j,k})$ is an upper triangular matrix. The eigenvalues of A are the principal diagonals of R.

Definition 2.2.1. The Rayleigh quotient $R_H(x)$ of a matrix H is defined as

$$R_H(x) = \frac{x^* H x}{x^* x}. \tag{2.2}$$

Lemma 2.2.2. *Given a Hermitian matrix H, let η_M and η_m be the maximum and minimum eigenvalues of H, respectively. Then*

$$\eta_m = \min_x R_H(x) = \min_{z:z^*z=1} z^* H z \tag{2.3}$$

$$\eta_M = \max_x R_H(x) = \max_{z:z^*z=1} z^* H z. \tag{2.4}$$

Proof. Suppose that e_m and e_M are eigenvectors corresponding to the minimum and maximum eigenvalues η_m and η_M, respectively. Then $R_H(e_m) = \eta_m$ and $R_H(e_M) = \eta_M$ and therefore

$$\eta_m \geq \min_x R_H(x) \tag{2.5}$$

$$\eta_M \leq \max_x R_H(x) \tag{2.6}$$

Since H is Hermitian we can write $H = UAU^*$, where U is unitary and

A is the diagonal matrix of the eigenvalues η_k, and therefore

$$\frac{x^*Hx}{x^*x} = \frac{x^*UAU^*x}{x^*x} = \frac{y^*Ay}{y^*y} = \frac{\sum_{k=1}^{n} |\gamma_k|^2 \eta_k}{\sum_{k=1}^{n} |\gamma_k|^2} \tag{2.7}$$

where $y = U^*x$ and we have taken advantage of the fact that U is unitary so that $x^*x = y^*y$. But for all vectors y, this ratio is bounded below by η_m and above by η_M and hence for all vectors x

$$\eta_m \leq R_H(x) \leq \eta_M \tag{2.8}$$

□

Lemma 2.2.3. *Let A be a matrix with eigenvalues α_k. Define the eigenvalues of the Hermitian nonnegative definite matrix A^*A to be $\lambda_k \geq 0$. Then*

$$\sum_{k=0}^{n-1} \lambda_k \geq \sum_{k=0}^{n-1} |\alpha_k|^2, \tag{2.9}$$

with equality if and only if A is normal.

Proof. The trace of a matrix is the sum of the diagonal elements of a matrix. The trace is invariant to unitary operations so that it also is equal to the sum of the eigenvalues of a matrix, i.e.,

$$\mathrm{Tr}(A^*A) = \sum_{k=0}^{n-1} (A^*A)_{k,k} = \sum_{k=0}^{n-1} \lambda_k. \tag{2.10}$$

Now, $A = URU^*$ and hence

$$\mathrm{Tr}(A^*A) = \mathrm{Tr}(R^*R) = \sum_{k=0}^{n-1}\sum_{j=0}^{n-1}|r_{j,k}|^2 = \sum_{k=0}^{n-1}|\alpha_k|^2 + \sum_{k\neq j}|r_{j,k}|^2 \geq \sum_{k=0}^{n-1}|\alpha_k|^2. \tag{2.11}$$

This equation will hold with equality if and only if R is diagonal and hence if and only if A is normal. □

2.3 Asymptotically Equivalent Sequences of Matrices

Definition 2.3.1. The Hilbert-Schmidt norm of an $n \times n$ matrix $A = [a_{k,j}]$ is defined by

$$\|A\| = \left(\frac{1}{n}\sum_{k=0}^{n-1}\sum_{j=0}^{n-1}|a_{k,j}|^2\right)^{1/2} \tag{2.12}$$

Definition 2.3.2. Two sequences of $n \times n$ matrices $\{A_n\}$ and $\{B_n\}$ are said to be asymptotically equivalent if

1. A_n and B_n are uniformly bounded, that is, $\|A_n\|, \|B_n\| \leq M < \infty, \forall n$

2. $A_n - B_n = D_n$ goes to zero as n tends to infinity, i.e., $\lim_{n\to\infty}\|A_n - B_n\| = 0$

Lemma 2.3.3. *Given two matrices A and B with eigenvalues $\{\alpha_k\}$ and $\{\beta_k\}$, respectively, we have*

$$\left|\frac{1}{n}\sum_{k=0}^{n-1}\alpha_k - \frac{1}{n}\sum_{k=0}^{n-1}\beta_k\right| \leq \|A - B\|. \tag{2.13}$$

Proof. Define the difference matrix $D = A - B = \{d_{k,j}\}$ so that

$$\sum_{k=0}^{n-1} \alpha_k - \sum_{k=0}^{n-1} \beta_k = \mathrm{Tr}(A) - \mathrm{Tr}(B) = \mathrm{Tr}(D). \tag{2.14}$$

Applying the Cauchy-Schwarz inequality to $\mathrm{Tr}(D)$ yields

$$|\mathrm{Tr}(D)|^2 = \left(\sum_{k=0}^{n-1} d_{k,k}\right)^2 \le n \sum_{k=0}^{n-1} |d_{k,k}|^2 \le n \sum_{k=0}^{n-1} \sum_{j=0}^{n-1} |d_{k,j}|^2 = n^2 \|D\|^2. \tag{2.15}$$

Taking the square root and dividing by n proves the lemma. □

Corollary 2.3.4. *Given two sequences of asymptotically equivalent matrices $\{A_n\}$ and $\{B_n\}$ with eigenvalues $\{\alpha_{n,k}\}$ and $\{\beta_{n,k}\}$, respectively, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} (\alpha_{n,k} - \beta_{n,k}) = 0, \tag{2.16}$$

and hence if either limit exists individually, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \alpha_{n,k} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \beta_{n,k}. \tag{2.17}$$

Proof. Let $D_n = \{d_{k,j}\} = A_n - B_n$. Then from Equation (2.17)

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{Tr}(D_n) = 0. \tag{2.18}$$

Dividing by n^2 and taking the limit results in

$$0 \le \left| \frac{1}{n} \mathrm{Tr}(D_n) \right|^2 \le \|D_n\|^2 \to 0 \text{ as } n \to \infty. \tag{2.19}$$

$\square$

Corollary 2.3.5. *Suppose that* $\{A_n\}$ *and* $\{B_n\}$ *are asymptotically equivalent sequences of matrices with eigenvalues* $\{\alpha_{n,k}\}$ *and* $\{\beta_{n,k}\}$*, respectively, and let* $f(x)$ *be any polynomial. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(f(\alpha_{n,k}) - f(\beta_{n,k})\right) = 0. \tag{2.20}$$

Hence if either limit exists separately then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\alpha_{n,k}) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\beta_{n,k}). \tag{2.21}$$

Theorem 2.3.6. *If* $F(x)$ *is a continuous complex function on* $[a, b]$*, there exist a sequence of polynomials* $p_n(x)$ *such that*

$$\lim_{n\to\infty} p_n(x) = F(x), \textit{uniformly on [a, b].} \tag{2.22}$$

Theorem 2.3.7. *Let* $\{A_n\}$ *and* $\{B_n\}$ *be asymptotically equivalent sequences of Hermitian matrices with eigenvalues* $\{\alpha_{n,k}\}$ *and* $\{\beta_{n,k}\}$*, respectively. Suppose there exist finite numbers* m *and* M *such that*

$$m \leq \alpha_{n,k}, \beta_{n,k} \leq M, \quad n = 1, 2, \ldots, \quad k = 0, 1, \ldots, n-1. \tag{2.23}$$

Let $F(x)$ *be an arbitrary function continuous on* $[m, M]$*. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \left(F(\alpha_{n,k}) - F(\beta_{n,k})\right) = 0, \tag{2.24}$$

and hence if either of the limits exists separately, then

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} F(\alpha_{n,k}) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} F(\beta_{n,k}). \tag{2.25}$$

Corollary 2.3.8. *Let $\{A_n\}$ and $\{B_n\}$ be asymptotically equivalent sequences of Hermitian matrices with eigenvalues $\{\alpha_{n,k}\}$ and $\{\beta_{n,k}\}$, respectively, such that $\alpha_{n,k}, \beta_{n,k} \geq m > 0$. If either limit exists, then*

$$\lim_{n\to\infty} (\det A_n)^{\frac{1}{n}} = \lim_{n\to\infty} (\det B_n)^{\frac{1}{n}}. \tag{2.26}$$

Proof. From Theorem 2.3.7 for $F(x) = \ln x$ we have,

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \ln \alpha_{n,k} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \ln \beta_{n,k} \tag{2.27}$$

and hence,

$$\lim_{n\to\infty} \exp\left[\frac{1}{n}\ln\left(\prod_{k=0}^{n-1} \alpha_{n,k}\right)\right] = \lim_{n\to\infty} \exp\left[\frac{1}{n}\ln\left(\prod_{k=0}^{n-1} \beta_{n,k}\right)\right]. \tag{2.28}$$

Equivalently,

$$\lim_{n\to\infty} \exp\left[\frac{1}{n}\ln(\det A_n)\right] = \lim_{n\to\infty} \exp\left[\frac{1}{n}\ln(\det B_n)\right]. \tag{2.29}$$

□

Chapter 3

Toeplitz Matrices and Their Properties

3.1 Introduction

Definition 3.1.1. A Toeplitz matrix is an $n \times n$ matrix $T_n = [t_{k,j}; k, j = 0, 1, \cdots, n-1]$ where $t_{k,j} = t_{k-j}$, that is, a matrix of the form

$$T_n = \begin{bmatrix} t_0 & t_{-1} & t_{-2} & \cdots & t_{-(n-1)} \\ t_1 & t_0 & t_{-1} & \cdots & t_{-(n-2)} \\ t_2 & t_1 & t_0 & \cdots & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & t_{n-3} & \cdots & t_0 \end{bmatrix}.$$

A great deal is known about the behavior of Toeplitz matrices. In this chapter, we deal with a sequence of Toeplitz matrices $\{T_n\}$ and the behavior of their eigenvalues as n goes to infinity.

Definition 3.1.2. The Fourier series representation of a complex-valued function $f(\lambda)$ can be written as:

$$f(\lambda) = \sum_{k=-\infty}^{\infty} t_k e^{ik\lambda} \tag{3.1}$$

where t_k are the complex Fourier coefficients given by:

$$t_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\lambda) e^{-ik\lambda}\, d\lambda. \tag{3.2}$$

Thus the sequence $\{t_k\}$ determines the function f and vice versa, hence the sequence of Toeplitz matrices is often denoted as $T_n(f)$.

3.2 Sequences of Toeplitz Matrices

Given the sums, product, eigenvalues, inverse, and determinants of circulant matrices, an obvious approach to the study of asymptotic properties of sequences of Toeplitz matrices is to approximate them by sequences asymptotically equivalent of circulant matrices and then applying the result developed so far.

Consider the infinite sequence $\{t_k\}$ and define the corresponding sequence of $n \times n$ Toeplitz matrices $T_n = [t_{k-j}; k, j = 0, 1, \ldots, n-1]$. Toeplitz matrices can be classified by the restrictions placed on the sequence t_k. The simplest class results if there is a finite m for which $t_k = 0$, $|k| > m$, in which case T_n is said to be a Banded Toeplitz matrix.

In the more general case where the t_k are not assumed to be zero for large k, there are two common constraints placed on the infinite sequence

$\{t_k; k = \ldots, -2, -1, 0, 1, 2, \ldots\}$ which defines all of the matrices T_n in the sequence. The most general is to assume that the t_k are square summable, i.e.,

$$\sum_{k=-\infty}^{\infty} |t_k|^2 < \infty. \tag{3.3}$$

We will make the stronger assumption that the t_k are absolutely summable, i.e.,

$$\sum_{k=-\infty}^{\infty} |t_k| < \infty. \tag{3.4}$$

Note that (3.4) is indeed a stronger constraint than (3.3) since

$$\left(\sum_{k=-\infty}^{\infty} |t_k|^2\right)^{1/2} \le \sum_{k=-\infty}^{\infty} |t_k|. \tag{3.5}$$

The assumption of absolute summability greatly simplifies the mathematics but does not alter the fundamental concepts of Toeplitz and circulant matrices involved. As the main purpose here is tutorial and we wish chiefly to relay the flavor and an intuitive feel for the results, we will confine interest to the absolutely summable case. The main advantage of (3.4) over (3.3) is that it ensures the existence and of the Fourier series $f(\lambda)$ defined by

$$f(\lambda) = \sum_{k=-\infty}^{\infty} t_k e^{ik\lambda} = \lim_{n\to\infty} \sum_{k=-n}^{n} t_k e^{ik\lambda}. \tag{3.6}$$

Not only does the limit in (3.9) converge if (3.4) holds, but it also converges uniformly for all λ, that is, we have

$$
\begin{aligned}
\left| f(\lambda) - \sum_{k=-n}^{n} t_k e^{ik\lambda} \right| &= \left| \sum_{k=-\infty}^{-n-1} t_k e^{ik\lambda} + \sum_{k=n+1}^{\infty} t_k e^{ik\lambda} \right| \\
&\leq \left| \sum_{k=-\infty}^{-n-1} t_k e^{ik\lambda} \right| + \left| \sum_{k=n+1}^{\infty} t_k e^{ik\lambda} \right| \\
&\leq \sum_{k=-\infty}^{-n-1} |t_k| + \sum_{k=n+1}^{\infty} |t_k|.
\end{aligned}
$$

where the right-hand side does not depend on λ, and it goes to zero as $n \to \infty$ from (3.4). Thus given $\epsilon > 0$, there is a single N not depending on λ, such that

$$
|f(\lambda) - \sum_{k=-n}^{n} t_k e^{ik\lambda}| \leq \epsilon, \quad \text{for all } \lambda \in [0, 2\pi], \text{ if } n \geq N. \tag{3.7}
$$

Furthermore, if (3.4) holds, then $f(\lambda)$ is Riemann integrable, and the t_k can be recovered from f from the ordinary Fourier inversion formula:

$$
t_k = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) e^{-ik\lambda} d\lambda. \tag{3.8}
$$

As a final useful property of this case, $f(\lambda)$ is a continuous function of $\lambda \in [0, 2\pi]$ except possibly at a countable number of points.

Definition 3.2.1. The essential supremum $M_f = esssup(f)$ of a real valued function f is the smallest number a for which $f(x) \leq a$ except on a set of measure 0.

Similarly, the essential infimum $m_f = essinf(f)$ is the largest number a such that $f(x) \geq a$ except on a set of measure 0.

The key idea here is to view M_f and m_f as the maximum and minimum values of f.

Lemma 3.2.2. *Let $\tau_{n,k}$ be the eigenvalues of a Toeplitz matrix $T_n(f)$. If $T_n(f)$ is Hermition. Then*

$m_f \leq \tau_{n,k} \leq M_f$.

Whether or not $T_n(f)$ is Hermition, we have, $\|T_n(f)\| \leq 2M_{|f|}$.

So that the sequence of Toeplitz matrices $Tn(f)$ is uniformly bounded over n if the essential supremum of $|f|$ is finite.

3.3 Banded Toeplitz matrices

Let T_n be a sequence of banded Toeplitz matrices of order $m+1$, that is, $t_i = 0$ unless $|i| \leq m$. Since we are interested in the behavior of T_n for large n, we choose $n \gg m$. As is easily seen from defination (3.1.1), T_n looks like a circulant matrix except for the upper left and lower right-hand corners; that is, each row is the row above shifted to the right one place. We can make a banded Toeplitz matrix exactly into a circulant if we fill in the upper right and lower left corners with the appropriate entries.

If a Toeplitz matrix is specified by a function f and hence denoted by $T_n(f)$, and the circulant matrix correspoding to it is denoted by $C_n(f)$. Then we have the following result.

Lemma 3.3.1. *The matrices T_n and C_n are asymptotically equivalent, i.e. both are bounded and* $\lim_{n\to\infty} |T_n - C_n| = 0$.

Theorem 3.3.2. *If $T_n(f)$ is a banded Toeplitz matrix with eigenvalues $\tau_{n,k}$, then for any positive integer s, we have*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \tau_{n,k}^s = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda)^s \, d\lambda$$

Furthermore, if f is real, then for any function $F(x)$ continuous on $[m_f, M_f]$, then we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} F(\tau_{n,k}) = \frac{1}{2\pi} \int_0^{2\pi} F(f(\lambda)) d\lambda.$$

3.4 Wiener Class Toeplitz Matrices

Definition 3.4.1. A sequence of Toeplitz matrices $T_n = [t_{k-j}]$ for which the t_k are absolutely summable is said to be in Wiener class. Similarly, a function $f(\lambda)$ defined on $[0, 2\pi]$ is said to be in the Wiener class if it has a Fourier series with absolutely summable Fourier coefficients.

Consider the case of f in the Wiener class, i.e., the case where the sequence $\{t_k\}$ is absolutely summable. As in the case of sequences of banded Toeplitz matrices, the basic approach is to find a sequence of circulant matrices $C_n(f)$ that is asymptotically equivalent to the sequence of Toeplitz matrices $T_n(f)$.

In the more general case under consideration, the construction of $C_n(f)$ is necessarily more complicated. Obviously, the choice of an appropriate sequence of circulant matrices to approximate a sequence of Toeplitz matrices is not unique, so we are free to choose a construction with the most desirable properties. It will, in fact, prove useful to consider two slightly different circulant approximations.

Since f is assumed to be in the Wiener class, we have the Fourier series representation:

$$f(\lambda) = \sum_{k=-\infty}^{\infty} t_k e^{ik\lambda} \tag{3.9}$$

where,

$$t_k = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) e^{-ik\lambda} d\lambda. \tag{3.10}$$

Define $C_n(f)$ to be the circulant matrix with the top row $(c_0^{(n)}, c_1^{(n)}, \ldots, c_{n-1}^{(n)})$ where

$$c_k^{(n)} = \frac{1}{n} \sum_{j=0}^{n-1} f\left(\frac{2\pi j}{n}\right) e^{2\pi ijk/n}. \tag{3.11}$$

Since $f(\lambda)$ is Riemann integrable, we have that for fixed k:

$$\lim_{n\to\infty} c_k^{(n)} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f\left(\frac{2\pi j}{n}\right) e^{2\pi ijk/n} = \frac{1}{2\pi} \int_0^{2\pi} f(\lambda) e^{ik\lambda} d\lambda = t_{-k} \tag{3.12}$$

and hence the $c_k^{(n)}$ are simply the sum approximations to the Riemann integrals giving t_{-k}. Now,

$$\begin{aligned}
\psi_{n,m} &= \sum_{k=0}^{n-1} c_k^{(n)} e^{-2\pi imk/n} \\
&= \sum_{k=0}^{n-1} \left(\frac{1}{n} \sum_{j=0}^{n-1} f\left(\frac{2\pi j}{n}\right) e^{2\pi ijk/n} \right) e^{-2\pi imk/n} \\
&= \sum_{j=0}^{n-1} f\left(\frac{2\pi j}{n}\right) \left(\frac{1}{n} \sum_{k=0}^{n-1} e^{-2\pi ik(j-m)/n} \right) \\
&= f\left(\frac{2\pi m}{n}\right).
\end{aligned} \tag{3.13}$$

Equation (3.13) in turn defines $C_n(f)$ since, if we are told that $C_n(f)$ is a circulant matrix with eigenvalues $f(2\pi m/n)$, $m = 0, 1, \ldots, n-1$, then

$$c_k^{(n)} = \frac{1}{n}\sum_{m=0}^{n-1} \psi_{n,m} e^{2\pi imk/n} = \frac{1}{n}\sum_{m=0}^{n-1} f\left(\frac{2\pi m}{n}\right) e^{2\pi imk/n}, \tag{3.14}$$

as in (3.11). Thus, either (3.11) or (3.14) can be used to define $C_n(f)$.

Theorem 3.4.2. *Let $T_n(f)$ be a sequence of Toeplitz matrices such that $f(\lambda)$ is in the Wiener class or, equivalently, that $\{t_k\}$ are absolutely summable. Let $\tau_{n,k}$ be the eigenvalues of $T_n(f)$, and let s be any positive integer. Then*

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} \tau_{n,k}^s = \frac{1}{2\pi}\int_0^{2\pi} f(\lambda)^s \, d\lambda.$$

Furthermore, if $f(\lambda)$ is real or, equivalently, the matrices $T_n(f)$ are Hermitian, then for any function $F(x)$ continuous on $[m_f, M_f]$, we have

$$\lim_{n\to\infty} \sum_{k=0}^{n-1} F(\tau_{n,k}) = \frac{1}{2\pi}\int_0^{2\pi} F(f(\lambda)) \, d\lambda.$$

3.5 Toeplitz Determinants

Suppose that $T_n(f)$ is a sequence of Hermitian Toeplitz matrices such that $f(\lambda) \geq m_f > 0$. Let $C_n(f)$ denote the sequence of circulant matrices constructed from f as (3.11). Then from (3.14) the eigenvalues of $C_n(f)$ are $f\left(\frac{2\pi m}{n}\right)$ for $m = 0, 1, \ldots, n-1$, and hence

$$\det(C_n(f)) = \prod_{m=0}^{n-1} f\left(\frac{2\pi m}{n}\right). \tag{3.15}$$

This in turn implies that

$$\ln(\det(C_n(f)))^{\frac{1}{n}} = \frac{1}{n} \ln \det(C_n(f)) = \frac{1}{n} \sum_{m=0}^{n-1} \ln f\left(2\pi \frac{m}{n}\right). \tag{3.16}$$

These sums are Riemann approximations to the limiting integral, leading to

$$\lim_{n\to\infty} \ln(\det(C_n(f)))^{\frac{1}{n}} = \int_0^1 \ln f(2\pi\lambda)\, d\lambda. \tag{3.17}$$

Exponentiating, using the continuity of the logarithm for strictly positive arguments, and changing the variables of integration yields

$$\lim_{n\to\infty} (\det(C_n(f)))^{\frac{1}{n}} = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \ln f(\lambda)\, d\lambda\right). \tag{3.18}$$

The asymptotic equivalence of $C_n(f)$ and $T_n(f)$, yields the following theorem:

Theorem 3.5.1. *Let $T_n(f)$ be a sequence of Hermitian Toeplitz matrices in the Wiener class such that $\ln f(\lambda)$ is Riemann integrable and $f(\lambda) \geq m_f$. Then*

$$\lim_{n\to\infty} (\det(T_n(f)))^{\frac{1}{n}} = \exp\left(\frac{1}{2\pi}\int_0^{2\pi} \ln f(\lambda)\, d\lambda\right). \tag{3.19}$$

Bibliography

[1] Philip J. Davis. *Circulant Matrices*. Wiley, 1979.

[2] Robert M. Gray. Toeplitz and circulant matrices: A review. *Foundations and Trends in Communications and Information Theory*, 2(3):155–239, 2006.

[3] Irwin Kra and Santiago R. Simanca. On circulant matrices. *Notices of the AMS*, 59(3):368–377, 2012.