A Generalization of Circulant Matrices for Non-Abelian Groups

Kent E. Morrison
Department of Mathematics
California Polytechnic State University
San Luis Obispo, CA 93407

kmorriso@calpoly.edu

August 18, 1998

Abstract

A circulant matrix of order n is the matrix of convolution by a fixed element of the group algebra of the cyclic group $\mathbf{Z}_n$. Replacing $\mathbf{Z}_n$ by an arbitrary finite group G gives the class of matrices that we call G-circulant. We determine the eigenvalues of such matrices with the tools of representation theory and the non-abelian Fourier transform.

Definition 1 *An n by n matrix C is* ***circulant*** *if there exist $c_0, \ldots, c_{n-1}$ such that the i, j entry of C is $c_{i-j \bmod n}$, where the rows and columns are numbered from 0 to $n-1$ and $k \bmod n$ means the number between 0 and $n-1$ that is congruent to k modulo n.*

For $n = 5$ a circulant matrix looks like

$$\begin{bmatrix} c_0 & c_4 & c_3 & c_2 & c_1 \\ c_1 & c_0 & c_4 & c_3 & c_2 \\ c_2 & c_1 & c_0 & c_4 & c_3 \\ c_3 & c_2 & c_1 & c_0 & c_4 \\ c_4 & c_3 & c_2 & c_1 & c_0 \end{bmatrix}$$

Definition 2 *Let $G = \{\sigma_1, \ldots, \sigma_n\}$ be a finite group of order n. An n by n matrix C is* ***G-circulant*** *(with respect to the ordering of G) if the entry in row i and column j is a function of $\sigma_i \sigma_j^{-1}$.*

A circulant matrix is a $\mathbf{Z}_n$-circulant matrix with the ordering $\mathbf{Z}_n = \{0, 1, \ldots, n-1\}$. We call a matrix **group-circulant** if it is G-circulant for some group G and an ordering of the elements of G.

Group-circulant matrices naturally arise as the transition matrices of Markov chains on finite groups. The state space is G and the probability of moving from τ to $\sigma\tau$ is p_σ.

At each step the current state is multiplied on the left by the element of G drawn from the probability distribution given by p. The transition matrix has in row σ and column τ the probability of moving from state τ to σ, which is $p_{\sigma\tau^{-1}}$. As an example, let G be the symmetric group S_n and define p to be concentrated uniformly on the transpositions. Since the transpositions generate S_n this Markov chain will tend to the uniform distribution on the entire group. If you were to use several random transpositions in an attempt to construct a random permutation, then you would like to know how quickly the approach to uniformity takes place. Such information can be extracted from the eigenvalues of the transition matrix.

Define the **group algebra** $\mathbf{C}[G]$ to be the set of functions $\phi : G \to \mathbf{C}$ with the usual operations of addition and scalar multiplication and with multiplication defined by

$$(\phi * \psi)(\sigma) = \sum_{\tau \in G} \phi(\sigma\tau^{-1})\psi(\tau).$$

It is, perhaps, simpler to define multiplication using the basis δ_σ, $\sigma \in G$,

$$\delta_\sigma(\tau) = \begin{cases} 1 & \sigma = \tau \\ 0 & \sigma \neq \tau \end{cases}$$

Define multiplication on the basis elements by $\delta_\sigma * \delta_\tau = \delta_{\sigma\tau}$ and extend by linearity. It is easy to verify that the two definitions are equivalent.

Theorem 3 *For ϕ in $\mathbf{C}[G]$ define the linear map*

$$C_\phi : \mathbf{C}[G] \to \mathbf{C}[G] : \psi \mapsto \phi * \psi.$$

Then the matrix of C_ϕ with respect to the basis $\{\delta_\sigma\}$ is G-circulant. Conversely, every G-circulant matrix arises in this way.

Proof Apply C_ϕ to the basis element δ_τ and extract the coefficient of δ_σ in the result.

$$\begin{aligned} C_\phi(\delta_\tau) &= \phi * \delta_\tau \\ &= \sum_{s \in G} \phi(s)\delta_s * \delta_\tau \\ &= \sum_{s \in G} \phi(s)\delta_{s\tau} \\ &= \sum_{\sigma \in G} \phi(\sigma\tau^{-1})\delta_\sigma \end{aligned}$$

Thus, the entry in row σ and column τ is $\phi(\sigma\tau^{-1})$. $\square$

The Fourier transform turns convolution, which is the multiplication in the group algebra, into multiplication and enables us to find the eigenvalues of C_ϕ. We summarize what is needed from the theory of representations of finite groups.

Definition 4 *For a finite group G let $\hat{G}$ denote the set of equivalence classes of irreducible representations of G. The set $\hat{G}$ is called the* **dual** *of G.*

It is convenient to pick representatives of the equivalence classes and to regard $\hat{G}$ as a set of specific representations. When G is abelian, $\hat{G}$ is also an abelian group and isomorphic to G, although not naturally isomorphic. When G is non-abelian, $\hat{G}$ does not have a group structure. Let $\hat{G} = \{\rho_1, \ldots, \rho_r\}$ and suppose that the dimension of ρ_i is d_i. It is useful to know that r, the number of irreducible representations, is also the number of conjugacy classes of G and that $\sum d_i^2 = n$.

Definition 5 *For $\phi \in \mathbf{C}[G]$ the **Fourier transform** of ϕ is the matrix valued function $\hat{\phi}$ on $\hat{G}$ defined by*

$$\hat{\phi}(\rho) = \sum_{s \in G} \phi(s)\rho(s).$$

Note that the sum above makes sense since all the matrices are the same size.

Theorem 6 (Fourier Inversion)

$$\phi(s) = \frac{1}{|G|} \sum_{\rho_i \in \hat{G}} d_i \mathrm{Tr}\,(\rho_i(s^{-1})\hat{\phi}(\rho_i)).$$

Theorem 7 *For ϕ and ψ in $\mathbf{C}[G]$,*

$$\widehat{\phi * \psi} = \hat{\phi}\hat{\psi},$$

where $(\hat{\phi}\hat{\psi})(\rho) = \hat{\phi}(\rho)\hat{\psi}(\rho)$ and the product is matrix multiplication.

Let $M_k(\mathbf{C})$ be the algebra of $k \times k$ complex matrices and define

$$\mathcal{M}[\hat{G}] := M_{d_1}(\mathbf{C}) \oplus \cdots \oplus M_{d_r}(\mathbf{C}).$$

Now $\mathbf{C}[G]$ and $\mathcal{M}[\hat{G}]$ both have dimension $n = |G|$. If we consider the Fourier transform of $\phi \in \mathbf{C}[G]$ as the k-tuple of matrices $(\hat{\phi}(\rho_1), \ldots, \hat{\phi}(\rho_r))$ or as an $n \times n$ matrix in block form

$$\hat{\phi}(\rho_1) \oplus \ldots \oplus \hat{\phi}(\rho_r),$$

then Fourier Inversion shows that the Fourier transform is a linear isomorphism. The previous theorem shows that it is also an algebra homomorphism where $\mathcal{M}[\hat{G}]$ has the product algebra structure. Therefore, the Fourier transform is an algebra isomorphism between $\mathbf{C}[G]$ and $\mathcal{M}[\hat{G}]$.

Note If G is abelian, then $d_i = 1$ and $r = n$, so that $\mathcal{M}[\hat{G}]$ can be identified with the algebra of complex valued functions on $\hat{G}$.

Theorem 8 *Let the eigenvalues of $\hat{\phi}(\rho_i)$ be $\lambda_{i,j}$, $1 \le j \le d_i$. Then these are the eigenvalues of C_ϕ and $\lambda_{i,j}$ has multiplicity d_i.*

Proof Let λ be an eigenvalue of C_ϕ with eigenvector ψ. Thus, $\phi * \psi = \lambda\psi$. Taking Fourier transforms we see that in $\mathcal{M}[\hat{G}]$

$$\hat{\phi}\hat{\psi} = \lambda\hat{\psi}.$$

Because of the block form for $\hat{\phi}$ the eigenvalues of C_ϕ are the union (over i) of the eigenvalues of multiplication by $\hat{\phi}(\rho_i)$ on $M_{d_i}(\mathbf{C})$. The eigenvalue equation

$$\hat{\phi}(\rho_i)\hat{\psi}(\rho_i) = \lambda\hat{\psi}(\rho_i)$$

is of the form $AB = \lambda B$ for square matrices A and B of size d_i. Hence each column of B is an eigenvector of A with eigenvalue λ. The action of A on the vector space of matrices $M_{d_i}(\mathbf{C})$ is equivalent to the direct sum of d_i copies of the action on $\mathbf{C}^{d_i}$ and so the multiplicity of the eigenvalue is multiplied by d_i. $\square$

Letting $G = \mathbf{Z}_n$, we can determine the well-known eigenvalue picture for circulant matrices. Let $\zeta = \exp(2\pi i/n)$. The irreducible representations are all one-dimensional and given by the characters $\rho_j(m) = \zeta^{jm}$. Thus, ρ_j maps the generator 1 in $\mathbf{Z}_n$ to ζ^j.

Theorem 9 (Diagonalization of Circulant Matrices) *Let C be the circulant matrix defined by $c_0, \ldots, c_{n-1}$ as in Definition 1. Then C is diagonalizable with eigenvalues $\lambda_0, \ldots, \lambda_{n-1}$ given by*

$$\lambda_j = \sum_{m=0}^{n-1} c_m \zeta^{jm}$$

and corresponding eigenvector

$$(1, \zeta^{-j}, \zeta^{-2j}, \ldots, \zeta^{-(n-1)j}).$$

Proof The matrix C is the matrix of convolution by ϕ where $\phi(m) = c_m$. By Theorem 8 the eigenvalues of C_ϕ are the eigenvalues of the 1×1 matrices $\hat{\phi}(\rho_j)$. Thus,

$$\lambda_j = \hat{\phi}(\rho_j) = \sum_m \phi(m)\rho_j(m) = \sum_m c_m \zeta^{jm}.$$

To get an eigenvector for λ_j let ψ_j be the element of the group algebra such that $\hat{\psi}_j = e_j := (0, \ldots, 0, 1, 0, \ldots, 0)$, where the e_j are the standard basis vectors of $\mathbf{C}^n$. Thus,

$$\hat{\psi}_j(\rho_k) = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$$

Note that $\hat{\phi}\hat{\psi}_j = \lambda_j \hat{\psi}_j$. Fourier Inversion gives

$$\begin{aligned} \psi_j(m) &= \frac{1}{n}\sum_k \rho_k(-m)\hat{\psi}_j(\rho_k) \\ &= \frac{1}{n}\zeta^{-jm} \end{aligned}$$

which is an eigenvector for convolution by ϕ in the group algebra. After multiplying by n we still have an eigenvector. Taking its coordinate representation with respect to the basis δ_m, $m = 0, \ldots, n-1$, gives the eigenvector in the statement of the theorem. $\square$

Corollary 10 *Let F be the $n \times n$ matrix with jk entry ζ^{-jk}, $0 \le j, k \le n-1$. Let C be a circulant matrix and let Λ be the diagonal matrix with diagonal entries λ_j. Then*

$$C = F\Lambda F^{-1}.$$

Furthermore, the jk entry of F^{-1} is

$$\frac{1}{n}\zeta^{jk}.$$

The center of the group algebra is the space of class functions. Recall that $\phi : G \to \mathbf{C}$ is a class function if $\phi(\tau\sigma\tau^{-1}) = \phi(\sigma)$ for all $\sigma, \tau \in G$. The center of $\mathcal{M}[\hat{G}]$ is the direct sum of the centers of the summands $M_{d_i}(\mathbf{C})$ and the center of the matrix algebra $M_{d_i}(\mathbf{C})$ is the space of scalar multiples of the identity. The Fourier transform maps the center of the group algebra onto the center of $\mathcal{M}[\hat{G}]$. Therefore, for a class function ϕ, the eigenvalues of C_ϕ are λ_i, $1 \le i \le r$, where λ_i corresponds to ρ_i and is the scalar such that

$$\hat{\phi}(\rho_i) = \lambda_i I_{d_i}.$$

With the trace we can isolate λ_i as

$$\lambda_i = \frac{1}{d_i} \sum_{\sigma \in G} \phi(\sigma) \operatorname{Tr} \rho_i(\sigma).$$

Let $G = S_3$ and order the elements

$$\iota, (1\,2), (2\,3), (1\,3), (1\,2\,3), (1\,3\,2).$$

Recall that there are three irreducible representations of S_3. Let ρ_1 be the trivial representation, ρ_2 the alternating representation,

$$\rho_2(\sigma) = (-1)^\sigma,$$

and let ρ_3 be the 2-dimensional representation

$$\iota \mapsto \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$(1\,2) \mapsto \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}, \quad (2\,3) \mapsto \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}, \quad (1\,3) \mapsto \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

$$(1\,2\,3) \mapsto \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, \quad (1\,3\,2) \mapsto \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}.$$

Let $\phi = \delta_{(1\,2\,3)}$. Using the convention that $\sigma\tau$ means σ followed by τ, i.e. $(1\,2)(2\,3) = (1\,3\,2)$, the matrix for C_ϕ is

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

The Fourier transform of ϕ is the element of $\mathcal{M}[\hat{S_3}] = \mathbf{C} \oplus \mathbf{C} \oplus M_2(\mathbf{C})$ given by

$$\begin{aligned} \hat{\phi}(\rho_1) &= 1 \\ \hat{\phi}(\rho_2) &= 1 \\ \hat{\phi}(\rho_3) &= \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}. \end{aligned}$$

The eigenvalues of $\hat{\phi}(\rho_3)$ are the cube roots of unity $-\frac{1}{2} \pm i\frac{\sqrt{3}}{2}$. Each of them has multiplicity two as an eigenvalue of C_ϕ. In addition, 1 is an eigenvalue of multiplicity 2. A quick check with MATLAB verifies that these are the eigenvalues of C_ϕ.