Kyungpook Math. J.
Volume 13, Number 1
June, 1973

INVERSES OF CIRCULANT MATRICES AND BLOCK CIRCULANT MATRICES

By George E. Trapp

1.0. Introduction.

Circulant matrices are part of the folklore. This paper has two purposes; first to summarize some of the elementary properties of circulants, and second to extend these results to block circulants. Circulants arise in the theory of statistical design [5], [6], [7] and in many applications in physics, see [1]. The books [3], [4], [9], [10], and [12] each contain some results on circulants.

We will consider $n\times n$ real matrices. Standard results of linear algebra will be assumed, see [3], or [8]. A matrix A is called a *circulant* whenever each row may be obtained from the row above by cyclically moving each element one place to the right. Example, the A given below is a circulant.

$$A = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$$

In section 2, we summarize some facts on nth roots of unity. It will be shown that the eigenvectors of circulants have as their components the nth roots. Section 3 deals with circulant properties basic to the study of block circulants. Section 4 considers block circulants. A by product of our approach to circulants is a very elementary algorithm to compute the Moore-Penrose generalized inverse of a symmetric circulant. In particular this gives a method of computing the true inverse of non-singular circulants.

In sectoin 5, we mention some related topics and future work. Because of our approach to block circulants, we are naturally lead to consider other block matrices. For example, the concept of a block Vandermonde matrix is discussed.

2.0. Roots of Unity.

Our approach is to diagonalize the circulant matrices. We therefore need to find the eigenvalues and eigenvectors of circulants.

To aid in our investigation of the eigenvalues and eigenvectors, we need to summarize some of the properties of the roots of unity. We will only sketch the proofs. Details may be found in [2] or [11].

The polynomial $X^n=1$ has n distinct solutions. $X=1$ is always a solution; for even n, $X=-1$ is another. In general the n roots of unity are given by $w_k=\cos\left(\frac{2\pi k}{n}\right)+i\sin\left(\frac{2\pi k}{n}\right)$ for $k=0, \cdots n-1$; notice $w_0=1$. Also it is easily seen that $w_k=w_1^k$. We will denote the complex conjugate of w by $\bar{w}$. For any polynomial equation we have that roots occur in complex conjugate pairs. Therefore $\bar{w}_k$ is also an nth root and can be written $\bar{w}_k=w_1^j$ for some j, $0\leq j\leq n-1$.

Since $w_k\bar{w}_k=1$, we see $w_1^{k+j}=1$. But w_1 is always a primitive nth root, and hence $n=k+j$. Therefore $\bar{w}_k=w_1^{n-k}$. Notice for $n=2p$, $w_p=-1$, and $\bar{w}_p=-1$ as expected.

Using the above result, the following orthogonality condition may be verified.

LEMMA 1. *Let* $\alpha_j=(1, w_j, w_j^2, \cdots, w_j^{n-1})$, *then* $\bar{\alpha}_k\cdot\alpha_j=n\delta_{kj}$ *where* δ_{ij} *the Kronecker delta* $\delta_{ij}=\begin{cases}1 & if\ i=j\\ 0 & if\ i\neq j.\end{cases}$

PROOF. $\bar{\alpha}_k\cdot\alpha_j=\sum_{t=0}^{n-1}(\bar{w}_k w_j)^t$. But $w_j=w_1^j$ and $\bar{w}_k=w_1^{n-k}$, therefore $\bar{\alpha}_k\cdot\alpha_j=\sum_{t=0}^{n-1}(w_1^{n+j-k})^t=\sum_{t=0}^{n-1}(w_1^{j-k})^t$ since $w_1^n=1$. Now $w_1^{j-k}\neq 1$ if $j\neq k$ and therefore w_1^{j-k} is one of the other nth roots. Since $X^n-1=(X-1)(X^{n-1}+X^n+\cdots+X+1)$ we see $\sum_{t=0}^{n-1}w^t=0$ for $w\neq 1$. If $j=k$ then $w_1^{j-k}=1$ and $\sum_{t=0}^{n-1}(w_1^{j-k})^t=\sum_{t=0}^{n-1}1^t=n$.

Lemma 1 will be used in determining the inverse of the matrix of eigenvectors of a circulant.

3.0 Diagonalization of Circulants.

In this section we show that every circulant may be diagonalized. Moreover, the matrix used in the diagonalization is the same for every circulant. This fact allows us to prove easily many results on circulants. If A is a circulant, then A has the form

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & \cdots & a_{n-2} \\ \vdots & & & \\ a_1 & a_2 & \cdots & a_0 \end{bmatrix}$$

Let $\alpha_j=(1, w_j, w_j^2, \cdots\cdots, w_j^{n-1})$ where w_j is an nth root of unity. Then direct computation yields

$$(1) \qquad A\alpha_j=\begin{bmatrix} a_0+a_1w_j+\cdots\cdots+a_{n-1}w_j^{n-1} \\ a_{n-1}+a_0w_j+\cdots+a_{n-2}w_j^{n-1} \\ \vdots \\ a_1+a_2w_j+\cdots\cdots+a_0w_j^{n-1} \end{bmatrix}.$$

Since $w_j^n=1$ we may rewrite the RHS of (1) as

$$(a_0+a_1w_j+\cdots\cdots+a_{n-1}w_j^{n-1})\begin{pmatrix} 1 \\ w_j \\ \vdots \\ w_j^{n-1} \end{pmatrix}.$$

Therefore α_j is an eigenvector of A with eigenvalue $\lambda_j=\sum_{k=0}^{n-1} a_k w_j^k$.

Let $P=[\alpha_0, \cdots\cdots, \alpha_{n-1}]$, here P is the matrix with columns the α_j. Then if D is the diagonal matrix of $\lambda_0\cdots\lambda_{n-1}$ we have $AP=PD$.

P is a Vandermonde matrix and since the w_j are distinct, it is non-singular. We then have

$$(2) \qquad P^{-1}AP=D.$$

Equation (2) may be written as $A=PDP^{-1}$. We will now show that by defining the matrix A as PDP^{-1}, for an arbitrary diagonal matrix D gives rise to a circulant. The following lemma supplies an explicit form for P^{-1}.

LEMMA 2. $P^{-1}=\frac{1}{n}P^*=\frac{1}{n}\bar{P}^t$.

PROOF. It is enough to show $\bar{P}^tP=nI$. $\bar{P}^tP$ may be written

$$\begin{pmatrix} \bar{\alpha}_0 \\ \vdots \\ \bar{\alpha}_{n-1} \end{pmatrix}[\alpha_0, \cdots\cdots, \alpha_{n-1}].$$

Therefore the ij element of $\bar{P}^tP$ is given by $\bar{\alpha}_i\cdot\alpha_j$. But Lemma 1 shows that $\bar{\alpha}_i\cdot\alpha_j=n\delta_{ij}$ as required.

We now proceed to show that if A is defined PDP^{-1}, then A is a circulant.

Let $D=\begin{pmatrix} d_0 & & 0 \\ & \ddots & \\ 0 & & d_{n-1} \end{pmatrix}$ be arbitrary. Then $PDP^{-1}=\frac{1}{n}PDP^*$ may be computed as follows:

$$PD=[\alpha_0,\cdots,\alpha_{n-1}]\begin{pmatrix} d_0 & & \\ & \ddots & \\ & & d_{n-1}\end{pmatrix}=[d_0\alpha_0,\ \cdots,\ d_{n-1}\alpha_{n-1}].$$

Then $\frac{1}{n}PDP^*=\frac{1}{n}[d_0\alpha_0,\ \cdots,\ d_{n-1}\alpha_{n-1}]\begin{pmatrix}\bar{\alpha}_0\\ \vdots \\ \bar{\alpha}_{n-1}\end{pmatrix}$

$$=\frac{1}{n}\begin{pmatrix}\sum d_i & \sum d_i\bar{w}_i & \cdots\cdots & \sum d_i(\bar{w}_i)^{n-1}\\ \sum d_i w_i & \sum d_i w_i\bar{w}_i & \cdots\cdots & \sum d_i w_i(\bar{w}_i)^{n-2}\\ \sum d_i w_i^{n-1} & \sum d_i w_i^{n-1}\bar{w}_i & \cdots\cdots & \sum d_i w_i^{n-1}(\bar{w}_i)^{n-1}\end{pmatrix}$$

Since $w_i\bar{w}_i=1$, and $w_i^k=\bar{w}_i^{n-k}$ we see that the resulting matrix is a circulant, with the a_i given by

$$a_i=\frac{1}{n}\sum_{j=0}^{n-1}d_j\bar{w}_j^i\ .$$

Summarizing the above we have the following theorem.

THOEREM 3. *With P as above, a matrix A is a circulant if and only if $\frac{1}{n}P^*AP$ is a diagonal matrix. Moreover, the eigenvalues of A are $\lambda_j=\sum_{k=0}^{n-1}a_k w_j^k$ with w_j an nth root of unity.*

The following theorem summarizes some of the major properties of circulants, the proof depends heavily on Theorem 3.

THEOREM 4. *Let A and B be circulants, then A^*, A^{-1} (if A is invertible) and $A+B$ are circulants. Moreover AB is a circulant and $AB=BA$.*

PROOF. A may be written $A=PDP^{-1}$, then if A is invertible $A^{-1}=PD^{-1}P^{-1}$ and by Theorem 3 A^{-1} is a circulant. Since $P^{-1}=\frac{1}{n}P^*$, $A^*=\frac{1}{n}PD^*P^*$ and then A^* is a circulant.

B may be written $B=PD_1P^{-1}$, therefore

$$\begin{aligned} A+B&=P(D+D_1)P^{-1} \text{ and}\\ AB&=P(DD_1)P^{-1}\\ &=P(D_1D)P^{-1}\\ &=BA\end{aligned}$$

and these are both circulants.

Given the representation $A=PDP^{-1}$ we may define $A^+=PD^+P^{-1}$, where D^+

is the diagonal matrix with diagonal $d_{ii}^{+}=1/d_{ii}$ if $d_{ii}\neq 0$ and $d_{ii}^{+}=0$ if $d_{ii}=0$. D^{+} is the Moore-Penrose Generalized inverse of D. The following facts follow easily

i) $AA^{+}A=(PDP^{-1})(PD^{+}P^{-1})(PDP^{-1})$
$=PDD^{+}DP^{-1}$
$=PDP^{-1}$ since $DD^{+}D=D$
$=A$

ii) $A^{+}AA^{+}=(PD^{+}P^{-1})(PDP^{-1})\ (PD^{+}P^{-1})$
$=A^{+}$ as above

iii) $AA^{+}=PDD^{+}P^{-1}=\frac{1}{n}PDD^{+}P^{*}$

$(AA^{+})^{*}=\frac{1}{n}PDD^{+}P^{*}=AA^{+}$ if $D=D^{*}$

iv) $(A^{+}A)^{*}=A^{+}A$ as in iii)

The four conditions given above demonstrate that the A^{+} so defined is the Moore-Penrose Generalized inverse of A. Since $P^{-1}=\frac{1}{n}P^{*}$, we have an elementary method of computing A^{+}. If in fact A is invertible, the above procedure yields A^{-1}. The major portion of computation other than matrix multiplication consists in accurately finding the nth roots of unity. Two comments are in order here: 1) since A^{+} is in fact a circulant, one needs only compute the first row of A^{+} (computation of the other rows could be used as a check), and 2) we are most interested in the generalized inverse for symmetric matrices-the above analysis tacitly assumed that A is symmetric.

4.0 Block Circulants.

Having presented some of the major properties of circulants, we now turn to block circulants. There seems to be some ambiguity involved in the term block circulant.

We define A to be a block circulant matrix if it has the following form

$$A=\begin{bmatrix} A_0 & A_1 & \cdots & A_{m-1} \\ A_{m-1} & A_0 & \cdots & A_{m-2} \\ \vdots & & & \\ A_1 & A_2 & \cdots & A_0 \end{bmatrix}.$$

The A_i are $n\times n$ matrices. We are most interested in the case when the A_i are themselves circulants although this is not needed in all that follows. Notice, except in unusual cases, the matrix A itself is not a circulant. Example, the A

given below is block circulant, but not circulant.

$$\left[\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ \hline 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{array}\right]$$

We would like to extend the results of section 3 to block circulants.

We begin by illustrating the approach for a 3 block matrix.

Let $E=\begin{bmatrix} A & B & C \\ C & A & B \\ B & C & A \end{bmatrix}$ where A, B, C are $n\times n$ matrices.

Let $1=r_0$, r_1 and r_2 be the 3rd roots of unity. If u is any vector and $r=r_i$, we have the following sequence of computations.

$$\begin{pmatrix} A & B & C \\ C & A & B \\ B & C & A \end{pmatrix}\begin{pmatrix} u \\ ru \\ r^2u \end{pmatrix}=\begin{pmatrix} (A+Br+Cr^2)u \\ (C+Ar+Br^2)u \\ (B+Cr+Ar^2)u \end{pmatrix}$$

$$=\begin{pmatrix} (A+Br+Cr^2)u \\ (Cr^2+A+Br)ru \\ (Br+Cr^2+A)r^2u \end{pmatrix}$$

Now pick u so that $(A+Br+Cr^2)\underline{u}=\lambda\underline{u}$, then

$$\begin{pmatrix} A & B & C \\ C & A & B \\ B & C & A \end{pmatrix}\begin{pmatrix} u \\ ru \\ r^2u \end{pmatrix}=\begin{pmatrix} (A+Br+Cr^2)u \\ (A+Br+Cr^2)ru \\ (A+Br+Cr^2)r^2u \end{pmatrix}$$

$$=\begin{pmatrix} \lambda(u) \\ \lambda(ru) \\ \lambda(r^2u) \end{pmatrix}=\lambda\begin{pmatrix} u \\ ru \\ r^2u \end{pmatrix}$$

Therefore our original 3 block circulant E has $3n$ eigenvalues. We obtain n for each choice of r. If A, B, C are circulant, then the matrix $A+Br_i+Cr_i^2$ is a circulant for each root r_i, and the eigenvalues can be easily determined; the matrix of eigenvectors are also easily determined; let P be as above.

Let $Q=\begin{bmatrix} P & P & P \\ r_0P & r_1P & r_2P \\ r_2^2P & r_1^2P & r_2^2P \end{bmatrix}$ then if $D=\begin{bmatrix} \Lambda_0 & 0 & 0 \\ 0 & \Lambda_1 & 0 \\ 0 & 0 & \Lambda_2 \end{bmatrix}$

with Λ_i the diagonal matrix with the eigenvalues of $A+Br_i+Cr_i^2$ on the diagonal, we have $EQ=QD$.

To complete the analogy with section 3, we must compute Q^{-1} and show that by defining a matrix QDQ^{-1} we obtain a block circulant.

The matrix Q may be rewritten as

$$Q=\begin{bmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{bmatrix}\begin{bmatrix} I & I & I \\ r_0 I & r_1 I & r_2 I \\ r_0^2 I & r_1^2 I & r_2^2 I \end{bmatrix}$$

Since P is invertible, we need only show that the second matrix on the right hand side is invertible to show Q is invertible. We will in fact find the inverse for the second matrix.

Consider the following; we will omit r_0 since it is 1.

$$\begin{bmatrix} I & I & I \\ I & r_1 I & r_2 I \\ I & r_1^2 I & r_2^2 I \end{bmatrix}\begin{bmatrix} I & I & I \\ I & r_1 I & r_1^2 I \\ I & r_2 I & r_2^2 I \end{bmatrix}$$

$$=\begin{bmatrix} 3I & I(1+r_1+r_2) & I(1+r_1^2+r_2^2) \\ I(1+r_1+r_2) & 3I & \cdots \\ I(1+r_1^2+r_2^2) & \cdots & \cdots \end{bmatrix}$$

Since $1+r_1+r_2=0$, we see that

$$\begin{bmatrix} I & I & I \\ I & r_1 I & r_2 I \\ I & r_1^2 I & r_2^2 I \end{bmatrix}^{-1}=\frac{1}{3}\begin{bmatrix} I & I & I \\ I & r_1 I & r_2 I \\ I & r_2 I & r_2^2 I \end{bmatrix}.$$

Therefore, using the fact $P^{-1}=\frac{1}{n}P^*$, we may write

$$Q^{-1}=\frac{1}{3n}\begin{bmatrix} P^* & P^* & P^* \\ P^* & r_1 P^* & r_1^2 P^* \\ P^* & r_2 P^* & r_2^2 P^* \end{bmatrix}.$$

Notice again, since $P^*=\overline{P}^t$, we have determined a simple formula for Q^{-1}.

We also have shown that $Q^{-1}EQ=D$. Moreover, we can now show that if D is an arbitrary diagonal matrix, then by defining $A=QDQ^{-1}$ we obtain a block circulant matrix, where the blocks are themselves circulants.

Let $D=\begin{bmatrix} D_0 & 0 & 0 \\ 0 & D_1 & 0 \\ 0 & 0 & D_2 \end{bmatrix}$, then QDQ^{-1} may be computed as follows:

QD is given by

$$\begin{bmatrix} P & P & P \\ P & r_1P & r_2P \\ P & r_1^2P & r_2^2P \end{bmatrix} \begin{bmatrix} D_0 & & 0 \\ & D_1 & \\ 0 & & D_2 \end{bmatrix} = \begin{bmatrix} PD_0 & PD_1 & PD_2 \\ PD_0 & r_1PD_1 & r_2PD_2 \\ PD_0 & r_1^2PD_1 & r_2^1PD_2 \end{bmatrix}.$$

Now QDQ^{-1} is given by

$$\begin{bmatrix} PD_0 & PD_1 & PD_2 \\ PD_0 & r_1PD_1 & r_2PD_2 \\ PD_0 & r_1^2PD_1 & r_2^2PD_2 \end{bmatrix} \begin{bmatrix} P^* & P^* & P^* \\ P^* & \bar{r}_1P^* & \bar{r}_1^2P^* \\ P^* & \bar{r}_2P^* & \bar{r}_2^2P^* \end{bmatrix}$$

$$= \begin{bmatrix} P(D_0+D_1+D_2)P^* & P(D_0+\bar{r}_1D_1+\bar{r}_2D_2)P^* & P(D_0+\bar{r}_1^2D_1+\bar{r}_2^2D_2)P^* \\ \cdots & & \\ \cdots & & \end{bmatrix}.$$

Therefore, since $r_i\bar{r}_i=1$ and $r_i^3=1$, we have a block circulant matrix. Moreover, since the blocks are given by $P(D_0+\bar{r}_1^iD_1+\bar{r}_2^iD_2)P^*$, our results of section 3 show that the blocks are themselves circulants.

We now state the main theorem for the general case. We will omit the proof since it is just an extension of the 3 block case given above.

THEOREM 5. *Let* $A=\begin{bmatrix} A_0 & \cdots & A_{m-1} \\ A_{m-1} & A_0\cdots & A_{m-2} \\ A_1 & A_2\cdots & A_0 \end{bmatrix}$ *be a block circulant with each* A_i *an* $n\times n$ *circulant matrix. Let* P *be the matrix of nth roots of unity as before. Let* $r_0, \cdots\cdots, r_{m-1}$ *be the mth roots of unity. If Q is given by the following block matrix*

$$Q=\begin{bmatrix} P & P & \cdots & P \\ r_0P & r_1P & \cdots & r_{m-1}P \\ \vdots & & & \\ r_0^{m-1}P & r_1^{m-1}P & \cdots & r_{m-1}^{m-1}P \end{bmatrix},$$

we have $Q^{-1}AQ=D$ *with* D *a matrix of diagonal blocks* $D_1, \cdots\cdots, D_{m-1}$ *where each* D_i *is diagonal. The diagonal elements are given by the eigenvalues of the matrix* $\sum_{k=0}^{m-1} A_kr_i^k$. *Moreover, given any diagonal matrix D, then for* $A=QDQ^{-1}$, *we have A is block circulant with each block being a circulant matrix.*

We also notice that Q^{-1} is given by the following.

$$Q^{-1}=\frac{1}{nm}\begin{bmatrix} P^* & \bar{r}_0P^* & \cdots & \bar{r}_0P^* \\ P^* & \bar{r}_1P^* & \cdots & \bar{r}_1^{m-1}P^* \\ \vdots & & & \\ P^* & \bar{r}_{m-1}P^* & \cdots & \bar{r}_{m-1}^{m-1}P^* \end{bmatrix}.$$

Theorem 5 allows us to prove the analogous theorem to theorem 4 for block circulants. In particular we have the following theorem.

THEOREM 6. *If A is an invertible block circulant with circulant blocks, then A^{-1} is of the same form.*

Theorem 5 also supplies a simple algorithm for finding A^{+}, and in particular for A^{-1} when A is non-singular.

Clearly, the results in this section may be extended to block-block circulants. We define a matrix R by the following. $R_{ij}=l_{i-1}^{j-1}\,Q$ where Q is as above, and l_{i-1} is a kth root of unity. If F is a block-block circulant, with k blocks, each block being an m block circulant, we have $FQ=QD$ with D diagonal, and the rest of the theory would follow.

5.0 Related Work.

The matrix P of the nth roots of unity is a Vandermonde matrix. Because of the special form of P, we did not use any of the theory of Vandermonde matrices (except for the invertibility condition). Possibly by reformulating the problem, this extensive theory could be used.

The matrix Q used in the block circulant case is not a Vandermonde matrix; but we term it a block Vandermonde matrix. It would be interesting to investigate properties of block Vandermonde matrices. In particular, the evaluation of the determinant and the computation of the inverse. Traub's work on Vandermonde matrices [13] could possibly be extended to this case.

In the case of symmetric circulants, we know that the eigenvalues are real, and that the eigenvectors may be chosen to be real. In our general treatment, we did not find the real eigenvectors. It might help the computational procedures if the real eigenvectors were used. At present this is not known.

Two other approaches to circulants are also available: we could use permutation matrices [3]; the block circulant case would then introduce the idea of block permutations. Also William Anderson has noticed that circulants arise from the characters of particular groups.

West Virginia University
Morgantown, West Virginia
26506 U.S.A.

REFERENCES

[1] Berlin, T.H. and Kac, M., *The Spherical Model of a Ferromagnet*, Phys, Rev., 86, pp. 821—835, 1952.

[2] Birkhoff, G. and MacLane, S., *A Survey of Modern Algebra*, New York, MacMillan Co., 1953.

[3] Browne, E. T., *Introduction to the Theory of Determinants and Matrices*, Chapel Hill, The University of North Carolina Press, 1958.

[4] Ferrar, *Algebra*, Oxford, Oxford University Press, 1941.

[5] John, J. A., Private communication.

[6] John, J. A., *Cyclic Incomplete Block Designs*, J. Roy. Stat., Ser. B, 28, pp. 345—360, 1966.

[7] Kempthorne, O., *A Class of Experimental Designs Using Blocks of Two Plots*, Annals of Math. Stat., 24, pp. 76—84, 1953.

[8] Marcus, M. and Minc, H., *Introduction to Linear Algebra*, New York, MacMillan Co., 1965.

[9] Mishina, A.P. and Proskuryakov, I. Y., *Higher Algebra*, Oxford, Pergamon Press, 1965.

[10] Muir, T., *A Treatise on the Theory of Determinants*, New York, Dover Publications Inc., 1960.

[11] Nehari, Z., *Introduction to Complex Analysis*, Boston, Allyn and Bacon Co., 1961.

[12] Proskuryakov, I. V., *Collected Problems in Linear Algebra*, Gostekhizoat, 1957.

[13] Traub, J. F. *Associated polynomials and uniform methods for the solution of linear problems*, 8, Siam Review, pp. 277—301, 1966.