

Reduction of Functions of Some Partitioned Matrices

By Victor Lovass-Nagy and David L. Powers

Abstract. In the numerical analysis of physical problems, there often arise large matrices which exhibit certain kinds of block-symmetry when partitioned appropriately. In this article, the structures of the frequently-occurring hyper-circulant and hyper-Jacobi matrices are examined, and it is shown how the calculation of any analytic function of such matrices may be reduced to the calculation of functions of the submatrices. Examples drawn from current engineering literature are given as well as small illustrative examples. ■

1. Introduction. The coefficient matrices of many matrix equations (both algebraic and differential) which arise in engineering and chemistry may easily be partitioned into square submatrices of equal order. It is known that if the submatrices commute with each other, the solution of the given matrix equation can be obtained in terms of the submatrices [2], [8]. Little seems to be known, however, about the reduction of functions of partitioned matrices whose submatrices do not commute to functions of the submatrices. (Some theorems on the eigenvalues of such partitioned matrices are given by B. Friedman [3] and Emilie V. Haynsworth [4].)

In a recent article [1], S. Charmonman discussed an algorithm for inverting a matrix partitioned into four (not necessarily commutable) blocks

$$(*) \quad \begin{bmatrix} A & B \\ B & A \end{bmatrix}$$

which takes advantage of the block-symmetry. In this paper, his results will be generalized in two ways: to hyper-circulant and hyper-Jacobi matrices, both generalizations of (*); and to any analytic function rather than just the inverse.

2. Circulant Matrices. If a_0, a_1, \dots, a_{n-1} are scalars it can be shown that the circulant matrix

$$C = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{bmatrix}$$

can be represented in the form

$$C = \sum_{k=0}^{n-1} u_k u_k^* \lambda_k$$

where $u_k = (1/\sqrt{n})(1, \omega^k, \omega^{2k}, \dots, \omega^{(n-1)k})$ ($k = 0, 1, \dots, n-1$) is an eigencolumn

Received April 1, 1968.

of C corresponding to the eigenvalue $\lambda_k = a_0 + a_1\omega^k + a_2\omega^{2k} + \dots + a_{n-1}\omega^{(n-1)k}$ (An asterisk on a vector indicates the conjugate transpose.) Here $\omega = \cos(2\pi/n) + i \sin(2\pi/n)$.

The hyper-circulant matrix

$$C_H = \begin{bmatrix} A_0 & A_1 & A_2 & \cdots & A_{n-1} \\ A_{n-1} & A_0 & A_1 & \cdots & A_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ A_1 & A_2 & A_3 & \cdots & A_0 \end{bmatrix}$$

where the A_k are $r \times r$ matrices, not necessarily commutative, may be represented analogously. For then,

$$C_H = \sum_{k=0}^{n-1} (u_k u_k^*) \times \Lambda_k.$$

(Here \times denotes the Kronecker product.) As before, u_k is an eigencolumn of the simple circulant of order n . The $r \times r$ matrix Λ_k is defined by

$$\Lambda_k = A_0 + A_1\omega^k + A_2\omega^{2k} + \dots + A_{n-1}\omega^{(n-1)k}.$$

Since the matrices $u_k u_k^*$ are idempotent and orthogonal, any analytic function f of the matrix C_H may be written as

$$f(C_H) = \sum_{k=0}^{n-1} (u_k u_k^*) \times f(\Lambda_k)$$

if the functions $f(\Lambda_k)$ exist.

In particular,

$$C_H^{-1} = \sum_{k=0}^{n-1} (u_k u_k^*) \times \Lambda_k^{-1}.$$

By the use of this representation, the computation of a function of a hyper-circulant with n rows and n columns of $r \times r$ submatrices is reduced to the computation of functions of n $r \times r$ matrices.

Two simple hyper-circulants give an illustration and extend Charmonman's work.

If

$$C_H = \begin{bmatrix} A & B & C \\ C & A & B \\ B & C & A \end{bmatrix},$$

then

$$\begin{aligned} f(C_H) &= \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \times f(A + B + C) + \frac{1}{3} \begin{bmatrix} 1 & \omega^2 & \omega \\ \omega & 1 & \omega^2 \\ \omega^2 & \omega & 1 \end{bmatrix} \times f(A + \omega B + \omega^2 C) \\ &\quad + \frac{1}{3} \begin{bmatrix} 1 & \omega & \omega^2 \\ \omega^2 & 1 & \omega \\ \omega & \omega^2 & 1 \end{bmatrix} \times f(A + \omega^2 B + \omega C) \end{aligned}$$

where $\omega = \cos(2\pi/3) + i \sin(2\pi/3)$.

If $B = C$, the matrix C_H is block-symmetric. Then

$$C_H = \begin{bmatrix} A & B & B \\ B & A & B \\ B & B & A \end{bmatrix}$$

and

$$f(C_H) = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \times f(A + 2B) + \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \times f(A - B).$$

3. A Special Jacobi Matrix. The $n \times n$ matrix

$$J = \begin{bmatrix} a & b & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ b & a & b & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & b & a & b & 0 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & b & a & b \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & b & a \end{bmatrix}$$

may be represented in a form similar to the above. In particular, if

$$v_k = \left(\frac{2}{n+1}\right)^{1/2} \left(\sin \frac{k\pi}{n+1}, \sin \frac{2k\pi}{n+1}, \dots, \sin \frac{nk\pi}{n+1}\right)$$

and v_k' denotes the transpose of the column matrix v_k , then J may be written as

$$J = \sum_{k=1}^n v_k v_k' \lambda_k \quad \text{where } \lambda_k = a + \beta_k b,$$

and

$$\beta_k = 2 \cos(k\pi/(n+1)) \quad (k = 1, 2, \dots, n).$$

Analogously, the hyper-Jacobi matrix

$$J_H = \begin{bmatrix} A & B & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ B & A & B & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & B & A & B & 0 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & B & A & B \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & B & A \end{bmatrix}$$

where A and B are (not necessarily commutable) $r \times r$ matrices, may be represented as $J_H = \sum_{k=1}^n (v_k v_k') \times \Lambda_k$, where the matrix $\Lambda_k = A + \beta_k B$. As before, any analytic function f of J_H may be written as $f(J_H) = \sum_{k=1}^n (v_k v_k') \times f(\Lambda_k)$.

As an illustration consider the simple case below: let

$$J_H = \begin{bmatrix} A & B & 0 \\ B & A & B \\ 0 & B & A \end{bmatrix},$$

then

$$f(J_H) = \frac{1}{2} \begin{bmatrix} \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{\sqrt{2}} & \frac{1}{2} \end{bmatrix} \times f\left(A + \frac{2}{\sqrt{2}}B\right) + \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} \times f(A) \\ + \frac{1}{2} \begin{bmatrix} \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \\ -\frac{1}{\sqrt{2}} & 1 & -\frac{1}{\sqrt{2}} \\ \frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \end{bmatrix} \times f\left(A - \frac{2}{\sqrt{2}}B\right).$$

4. Examples. Three examples are presented which appear in engineering problems. The first (see [5, p. 101]) is a special hyper-circulant matrix

$$C_H = \begin{bmatrix} A & I & 0 & \cdots & 0 & 0 & I \\ I & A & I & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ I & 0 & 0 & \cdots & 0 & I & A \end{bmatrix}$$

in which $A_0 = A, A_1 = A_{n-1} = I$ and the rest of the $A_k = 0$. In order to invert C_H it is necessary to find and invert the matrices Λ_k . By the definition above,

$$\Lambda_k = A + I\omega^k + I\omega^{(n-1)k} = A + 2(\cos(2k\pi/n))I.$$

In the particular example cited, A also is a circulant of the same form as C_H :

$$\begin{bmatrix} -4 & 1 & 0 & \cdots & 0 & 0 & 1 \\ 1 & -4 & 1 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 1 & 0 & 0 & \cdots & 0 & 1 & -4 \end{bmatrix}.$$

Thus Λ_k is also a circulant and may be inverted by using the formulas in Section 2.

The second example is a special hyper-Jacobi matrix which occurred in calculations for the load capacity of a slider bearing with non-Newtonian lubricant [7]. If a symmetry condition on the center-line is not used, the matrix has the form

$$J_H = \begin{bmatrix} A & B & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ B & A & B & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & B & A & B & 0 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & B & A & B \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & B & A \end{bmatrix},$$

where

$$A = \begin{bmatrix} 1 & a_1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ c_1 & 1 & a_2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & c_2 & 1 & a_3 & 0 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \cdots & c_{n-2} & 1 & a_{n-1} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & c_{n-1} & 1 \end{bmatrix}$$

and $B = \text{diag}(b_1, b_2, \dots, b_n)$. Note that $AB \neq BA$. The large matrix J_H may be inverted by formula, involving the inverses of the smaller matrices $A + 2(\cos(k\pi/(n+1)))B$.

The third example features a matrix which is neither hyper-Jacobi nor hyper-circulant in structure, but which may be attacked by "cascading." In an article on polymer chains [6], there appears the matrix P partitioned into 3×3 submatrices:

$$P = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \\ A_2 & A_1 & A_4 & A_3 \\ A_3 & A_4 & A_1 & A_2 \\ A_4 & A_3 & A_2 & A_1 \end{bmatrix}, \quad A_i A_j \neq A_j A_i.$$

This matrix fails to have the structure of the hyper-circulant because A_2 and A_4 are out of place in the second and fourth rows. However, if the matrix is partitioned again in the form

$$P = \begin{bmatrix} M_1 & M_2 \\ M_2 & M_1 \end{bmatrix}$$

it does have the hyper-circulant form. Thus, letting $L_i = u_i u_i^*$, one has

$$P = L_1 \times (M_1 + M_2) + L_2 \times (M_1 - M_2)$$

where

$$L_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad L_2 = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

Now each of the matrices M_1 and M_2 is a partitioned matrix having the hyper-circulant form, so

$$\begin{aligned} M_1 + M_2 &= L_1 \times (A_1 + A_2 + A_3 + A_4) + L_2 \times (A_1 - A_2 + A_3 - A_4), \\ M_1 - M_2 &= L_1 \times (A_1 + A_2 - A_3 - A_4) + L_2 \times (A_1 - A_2 - A_3 + A_4). \end{aligned}$$

Finally we may write

$$\begin{aligned} P &= (L_1 \times L_1) \times (A_1 + A_2 + A_3 + A_4) + (L_1 \times L_2) \\ &\quad \times (A_1 - A_2 + A_3 - A_4) + (L_2 \times L_1) \times (A_1 + A_2 - A_3 - A_4) \\ &\quad + (L_2 \times L_2) \times (A_1 - A_2 - A_3 + A_4). \end{aligned}$$

Thus the determination of a function of P (which in the reference is a 12×12 matrix) is reduced to the determination of four functions of matrices of one fourth the original order.

It should be mentioned that the matrices $L_i \times L_j$ are idempotent and

$$(L_i \times L_j)(L_p \times L_q) = \delta_{ip} \delta_{jq}; \quad i = 1, 2; j = 1, 2; p = 1, 2; q = 1, 2.$$

$$\sum_{i=1}^2 \sum_{j=1}^2 L_i \times L_j = I_4.$$

Hence

$$\begin{aligned} P^{-1} &= (L_1 \times L_1) \times (A_1 + A_2 + A_3 + A_4)^{-1} + (L_1 \times L_2) \\ &\quad \times (A_1 - A_2 + A_3 - A_4)^{-1} + (L_2 \times L_1) \\ &\quad \times (A_1 + A_2 - A_3 - A_4)^{-1} + (L_2 \times L_2) \\ &\quad \times (A_1 - A_2 - A_3 + A_4)^{-1} \end{aligned}$$

and an analytic function of P can be written as

$$\begin{aligned}
 f(P) &= (L_1 \times L_1) \times f(A_1 + A_2 + A_3 + A_4) + (L_1 \times L_2) \\
 &\quad \times f(A_1 - A_2 + A_3 - A_4) + (L_2 \times L_1) \\
 &\quad \times f(A_1 + A_2 - A_3 - A_4) + (L_2 \times L_2) \\
 &\quad \times f(A_1 - A_2 - A_3 + A_4).
 \end{aligned}$$

Obviously,

$$\begin{aligned}
 L_1 \times L_1 &= \frac{1}{2} \begin{bmatrix} L_1 & L_1 \\ L_1 & L_1 \end{bmatrix}; & L_1 \times L_2 &= \frac{1}{2} \begin{bmatrix} L_2 & L_2 \\ L_2 & L_2 \end{bmatrix}; \\
 L_2 \times L_1 &= \frac{1}{2} \begin{bmatrix} L_1 & -L_1 \\ -L_1 & L_1 \end{bmatrix}; & L_2 \times L_2 &= \frac{1}{2} \begin{bmatrix} L_2 & -L_2 \\ -L_2 & L_2 \end{bmatrix}.
 \end{aligned}$$

5. Comment. If a partitioned matrix has commutative submatrices and the block-circulant or block-Jacobi structure, there are two ways to represent functions of that matrix by the aid of idempotent matrices. We show the difference between these methods with the following illustration.

Consider the partitioned matrix

$$M = \begin{bmatrix} A & B & B \\ B & A & B \\ B & B & A \end{bmatrix}$$

where A and B are 2×2 matrices.

I. *Method Based on Commutativity of Submatrices.* It is sufficient that A and B be simultaneously reducible to triangular form in order to find the eigenvalues of $f(M)$. (See [8, p. 588].) We shall assume in addition that A and B are polynomials in some matrix C . Say $A = \alpha_0 I + \alpha_1 C$, $B = \beta_0 I + \beta_1 C$. Then

$$M = \begin{bmatrix} \alpha_0 & \beta_0 & \beta_0 \\ \beta_0 & \alpha_0 & \beta_0 \\ \beta_0 & \beta_0 & \alpha_0 \end{bmatrix} \times I + \begin{bmatrix} \alpha_1 & \beta_1 & \beta_1 \\ \beta_1 & \alpha_1 & \beta_1 \\ \beta_1 & \beta_1 & \alpha_1 \end{bmatrix} \times C = E \times I + F \times C.$$

If C has the distinct eigenvalues λ_1 and λ_2 and the associated idempotent matrices $L_1(C)$ and $L_2(C)$, then $C = \lambda_1 L_1(C) + \lambda_2 L_2(C)$ and

$$M = (E + \lambda_1 F) \times L_1(C) + (E + \lambda_2 F) \times L_2(C)$$

and an analytic function $f(M)$ can be written

$$f(M) = f(E + \lambda_1 F) \times L_1(C) + f(E + \lambda_2 F) \times L_2(C),$$

where $L_1(C) = (C - \lambda_2 I)/(\lambda_1 - \lambda_2)$ and $L_2(C) = (C - \lambda_1 I)/(\lambda_2 - \lambda_1)$.

This method of representation is to some extent implied by the results of Williamson [8] and Egerváry [2].

II. *Hyper-circulant-Structure Method.* Whether or not $AB = BA$, the matrix M has the hyper-circulant (block-circulant) structure and so $M = P_1 \times (A + 2B) + P_2 \times (A - B)$, where P_1 and P_2 are the idempotent matrices associated with any 3×3 circulant matrix

$$P_1 = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}; \quad P_2 = \frac{1}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

and

$$f(M) = P_1 \times f(A + 2B) + P_2 \times f(A - B).$$

Note that

(1) In Method I, the idempotent matrices are on the right, in Method II, on the left.

(2) In Method I, the idempotent matrices are 2×2 matrices, in Method II, they are 3×3 matrices.

(3) In Method I, the function of the original 6×6 matrix M is reduced to functions of two 3×3 matrices, in Method II, the same function of M is reduced to functions of two 2×2 matrices.

(4) Method II sets neither commutativity requirement nor the requirement of being simultaneously-similar to triangular matrices on A and B .

6. Conclusion. Many matrix problems encountered in science and engineering involve hyper-circulant and hyper-Jacobi matrices discussed in this article. At the present time, iterative methods are widely used in their solution. Evidently the amount of time required for solution can be vastly reduced if inverses and other functions are calculated by the formulas presented here rather than by iteration. Moreover, accuracy is improved, and questions of stability or convergence do not even arise.

The results of Section 3 may be extended to hyper-Jacobi matrices in which $A + \alpha B$ and $A + \beta B$ appear instead of A in the upper left and lower right corners. Only the elements of the column matrices v_k and those of the $r \times r$ matrices Δ_k used in the formulas are changed.

The technique of representing a partitioned matrix in terms of the eigencolumns and eigenrows of an ordinary matrix may be extended to any partitioned matrix of the form $\sum_k f_k(Q) \times M_k = P$, where $f_k(Q)$ are all functions of a common $n \times n$ matrix Q and the M_k are arbitrary $r \times r$ matrices. Then P may be represented in terms of the eigenvectors of Q .

Department of Mathematics
Clarkson College of Technology
Potsdam, New York 13676

1. S. CHARMONMAN, "An efficient algorithm for inverting a block-symmetric matrix," *Math. Comp.*, v. 21, 1967, pp. 715-717.
2. E. EGERVÁRY, "On hypermatrices whose blocks are commutable in pairs and their applications in lattice-dynamics," *Acta Sci. Math. Szeged*, v. 15, 1954, pp. 211-222. MR 16, 327.
3. B. FRIEDMAN, "Eigenvalues of compound matrices," *Research Report TW-16*, Mathematics Research Group, New York University Washington Square College of Arts and Science, 1951.
4. E. V. HAYNSWORTH, "Special types of partitioned matrices," *J. Res. Nat. Bur. Standards Sect. B* 65B, 1961, pp. 7-12. MR 27 #165.
5. R. W. HOCKNEY, "A fast direct solution of Poisson's equation using Fourier analysis," *J. Assoc. Comput. Mach.*, v. 12, 1965, pp. 95-113.
6. E. W. MONTROLL, "Markoff chains and excluded volume effect in polymer chains," *J. Chem. Phys.*, v. 18, 1950, pp. 734-743. MR 12, 114.
7. F. E. STEIDLER & H. H. HOROVITZ, "The calculated load-carrying ability of non-Newtonian lubricants in hydro-dynamic bearings," *Chemical Engineering Progress Symposium Series*, no. 42, Vol. 59, 1963, pp. 99-107.
8. J. WILLIAMSON, "The latent root of a matrix of special type," *Bull. Amer. Math. Soc.*, v. 37, 1931, pp. 585-590.