MATRICES AND CONVOLUTIONS OF ARITHMETIC FUNCTIONS

E. E. GUERIN
Seton Hall University, South Orange, NJ 07079

1. INTRODUCTION

The purpose of this paper is to relate certain matrices with integer entries to convolutions of arithmetic functions.

Let n be a positive integer, let α, β, and γ be arithmetic functions (complex-valued functions with domain the set of positive integers), and let $\alpha_{[n]}$ denote the $1 \times n$ matrix $[\alpha(1)\ \alpha(2)\ \ldots\ \alpha(n)]$.

We define the $n \times n$ divisor matrix $D_n = (d_{ij})$ by $d_{ij} = 1$ if $i \mid j$, $d_{ij} = 0$ otherwise. Both D_n and its inverse, D_n^{-1}, are upper triangular matrices. The arithmetic functions ν_k, σ, and ε are defined by $\nu_k(n) = n^k$ for $k = 0, 1, 2$, $\sigma(n) = \sum_{d \mid n} d$, and $\varepsilon(n) = 1$ if $n = 1$, $\varepsilon(n) = 0$ if $n > 1$. We also consider the divisor function τ, the Moebius function μ, and Euler's ϕ-function. We observe that

$$\nu_{0[n]} D = \tau_{[n]}, \tag{1}$$

$$\nu_{1[n]} D = \sigma_{[n]}, \tag{2}$$

$$\varepsilon_{[n]} D_n^{-1} = \mu_{[n]}, \tag{3}$$

$$\nu_{1[n]} D_n^{-1} = \phi_{[n]}. \tag{4}$$

These matrix formulas, which can be used to evaluate arithmetic functions as in [2], are consequences of the following equations which involve the Dirichlet convolution, $*_D$.

$$\nu_0 *_D \nu_0 = \tau, \tag{1$'$}$$

$$\nu_1 *_D \nu_0 = \sigma, \tag{2$'$}$$

$$\varepsilon *_D \mu = \mu, \quad \varepsilon = \mu *_D \nu_0, \tag{3$'$}$$

$$\nu_1 *_D \mu = \phi, \quad \phi *_D \nu_0 = \nu_1. \tag{4$'$}$$

As an illustration, consider matrices D_6 and D_6^{-1} which appear below.

$$D_6 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 1 & 0 & 1 \\ & & 1 & 0 & 0 & 1 \\ & & & 1 & 0 & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{bmatrix}, \qquad D_6^{-1} = \begin{bmatrix} 1 & -1 & -1 & 0 & -1 & 1 \\ & 1 & 0 & -1 & 0 & -1 \\ & & 1 & 0 & 0 & -1 \\ & & & 1 & 0 & 0 \\ & & & & 1 & 0 \\ & & & & & 1 \end{bmatrix}.$$

Any omitted entry is assumed to be zero. By (2),

$$[1\ 2\ 3\ 4\ 5\ 6]D_6 = [\sigma(1)\ \sigma(2)\ \sigma(3)\ \sigma(4)\ \sigma(5)\ \sigma(6)],$$

so that $\sigma(6) = \sum_{d|6} d = \sum_{d|6} \nu_1(d) = (\nu_1 *_D \nu_0)(6)$. And by (4),

$$[1\ 2\ 3\ 4\ 5\ 6]D_6^{-1} = [\phi(1)\ \phi(2)\ \phi(3)\ \phi(4)\ \phi(5)\ \phi(6)],$$

so that $\phi(6) = 1 - 2 - 3 + 6 = (\nu_1 *_D \mu)(6)$.

These observations lead us to define and illustrate matrix-generated convolutions.

2. MATRIX-GENERATED CONVOLUTIONS

Suppose that $G = (g_{ij})$ is an infinite dimentional (0, 1)-matrix with $g_{ij} = 1$ if $i = j$ and $g_{ij} = 0$ if $i > j$, and that the 1's in column n of G appear in rows $n_1, n_2, \ldots, n_k$ ($n_1 < n_2 < \ldots < n_k = n$). We say that G generates the convolution $*_G$ defined by

$$(\alpha *_G \beta)(n) = \sum_{v=1}^{k} d(n_v)\beta(n_{k+1-v}),\ n = 1, 2, 3, \ldots .$$

Clearly, $*_G$ is a commutative operation on the set of arithmetic functions. We denote by G_n the $n \times n$ submatrix of $G = (g_{ij})$ with $1 \leqq i \leqq n$, $1 \leqq j \leqq n$.

The convolutions in Examples 1-4 below are defined and referenced in [3].

Example 1: The matrix $D = (d_{ij})$, with $d_{ij} = 1$ if $i|j$, $d_{ij} = 0$ otherwise, generates the Dirichlet convolution $*_D$. D_n is the $n \times n$ divisor matrix, and the set $\{n_1, n_2, \ldots, n_k\}$ is the set of positive divisors of n.

Example 2: The unitary convolution is generated by the matrix $U = (u_{ij})$ with $u_{ij} = 1$ if $i \leqq j$ and $i|j$ and i and j/i are relatively prime, $u_{ij} = 0$ otherwise.

Example 3: The matrix $C = (c_{ij})$ defined by $c_{ij} = 1$ if $i \leqq j$, $c_{ij} = 0$ otherwise, generates a convolution $*_C$ related to the Cauchy product. Since $\{n_1, n_2, \ldots, n_k\} = \{1, 2, \ldots, n\}$, we have

$$(\alpha *_C \beta)(n) = \alpha(1)\beta(n) + \alpha(2)\beta(n-1) + \cdots + \alpha(n)\beta(1).$$

Example 4: For a fixed prime p, let the matrix $L = (l_{ij})$ be defined by $l_{ij} = 1$ if $i \leqq j$ and $p \nmid \binom{j-1}{i-1}$, $l_{ij} = 0$ otherwise. The convolution $*_L$ generated by L is related to the Lucas product. The entries shown in the matrix L_{14} for $p = 3$ are easily determined by the use of a basis representation criterion given in [1].

$$
\begin{array}{c}
L_{14} = \\
(p = 3)
\end{array}
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\
 & & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
 & & & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\
 & & & & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\
 & & & & & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & 1 & 0 & 0 & 0 & 0 & 0 \\
 & & & & & & & & & 1 & 1 & 1 & 1 & 1 \\
 & & & & & & & & & & 1 & 1 & 0 & 1 \\
 & & & & & & & & & & & 1 & 0 & 0 \\
 & & & & & & & & & & & & 1 & 1 \\
 & & & & & & & & & & & & & 1
\end{bmatrix}
$$

3. A GENERAL MOEBIUS FUNCTION

In view of (3′), we next define a general Moebius function μ_G by $\nu_0 *_G \mu_G = \varepsilon$. It is immediate from $G_n^{-1}G_n = I_n$ (the $n \times n$ identity matrix) that

$$\text{if } G_n^{-1} = (\bar{g}_{ij}) \text{ then } \bar{g}_{ij} = \mu(j) \text{ for } j = 1, 2, \ldots, n \text{ and } n = 1, 2, 3, \ldots \tag{5}$$

For example, the elements in row one of D_6^{-1} are $\mu_D(1) = \mu(1), \mu(2), \ldots, \mu(6)$ (in that order). The values of the unitary, Cauchy, and Lucas Moebius functions given in [3] agree with corresponding entries in row one of U_n, C_n, and L_n, respectively. Property (5) implies $\varepsilon_{[n]}G_n^{-1} = \mu_{G[n]}$, which is a generalization of (3).

The following three properties are related to the Moebius function and are stated for future reference.

$$\alpha *_G \varepsilon = \alpha \text{ for all arithmetic functions } \alpha. \tag{6}$$

$$*_G \text{ is an associative operation on the set of arithmetic functions.} \tag{7}$$

$$\text{If } g_{ij} = 0 \text{ then } \bar{g}_{ij} = 0, \text{ where } G_n^{-1} = (\bar{g}_{ij}),\ n = 1, 2, 3, \ldots . \tag{8}$$

Property (6) is equivalent to

$$g_{1j} = 1 \text{ for } j = 1, 2, 3, \ldots . \tag{6′}$$

For (6′) clearly implies (6); and if $g_{1n} = 0$ for some n, and α is such that $\alpha(n) \neq 0$, then $(\alpha *_G \varepsilon)(n) = 0 \neq \alpha(n)$.

Example 5: Let the matrix $P = (p_{ij})$ be defined by $p_{ij} = 1$ if $i \leq j$ and i and j are of the same parity, $p_{ij} = 0$ otherwise. Evidently, (6′) and (6) do not hold here. For example, $(\nu_0 *_P \varepsilon)(2) = \nu_0(2)\varepsilon(2) = 0 \neq \nu_0(2)$. Although ε', defined by $\varepsilon'(1) = \varepsilon'(2) = 1$, $\varepsilon'(n) = 0$ if $n > 2$, satisfies $\alpha *_P \varepsilon' = \alpha$ for all arithmetic functions α, ε' is not related to matrix multiplication in $G_n^{-1}G_n = I_n$ in the desirable way that ε is.

We note that if (6) and (7) hold then we can apply Moebius inversion in the form $\alpha = \nu_0 *_G \beta$ iff $\beta = \mu_G *_G \alpha$ [as illustrated in (4′)]. It is clear that

(6) holds and well known that (7) holds for the convolutions in Examples 1-4; so (8) holds as well, as can be verified by direct computation or by application of the following theorem.

Theorem 1: Property (7) implies property (8).

Proof: Assume that (8) is false. Let j be the smallest positive integer such that for some i we have $g_{ij} = 0$ and $\overline{g}_{ij} \neq 0$; let this $j = n$. Consider the largest value of i such that $g_{in} = 0$ and $\overline{g}_{in} \neq 0$; let this $i = t$. It follows by the assumptions and $G_n G_n^{-1} = I_n$ that $g_{tt} = 1$, $g_{tn} = 0$, $\overline{g}_{tn} \neq 0$, there is an integer r such that $t < r < n$ and $g_{tr} = 1$, and $g_{rn} = 1$. Since $r \in \{n_1, \ldots, n_k\}$ and $g_{tr} = 1$, then $\alpha(t)$ is a factor in some term of

$$\big((\alpha *_G \beta) *_G \gamma\big)(n).$$

But no term of $\big(\alpha *_G (\beta *_G \gamma)\big)(n)$ has a factor $\alpha(t)$ because $t \notin \{n_1, \ldots, n_k\}$. Therefore, (7) is false and the proof is complete. ■

4. THE MAIN THEOREM

We now define some special functions and matrices leading to the main result in this paper. Assume that the matrix G generates the convolution $*_G$ and define the arithmetic functions A and B by

$$A(n) = \sum_{i=1}^{n} g_{in}\alpha(i) \quad \text{and} \quad B(n) = \sum_{i=1}^{n} \overline{g}_{in}\,\beta(i).$$

Then for $n = 1, 2, 3, \ldots$, we have

$$\alpha_{[n]} G_n = A_{[n]} \tag{9}$$

and

$$\beta_{[n]} G_n^{-1} = B_{[n]}. \tag{10}$$

Define $G_n^S = (s_{ij})$ to be the $n \times n$ matrix with $s_{ij} = 1$ if $i = n_v$ and $j = n_{k+1-v}$, $v = 1, 2, \ldots, k$, $s_{ij} = 0$ otherwise. Note that G_n^S is a symmetric (0, 1)-matrix with at most one nonzero entry in any row or column. If M^t denotes the transpose of a matrix M, then

$$(\alpha *_G \beta)(n) = \alpha_{[n]} G_n^S \big(\beta_{[n]}\big)^t \tag{11}$$

and

$$(A *_G B)(n) = A_{[n]} G_n^S \big(B_{[n]}\big)^t. \tag{12}$$

The matrix $G_n G_n^S$ is of special interest and can be characterized as follows.

Column n_v of $G_n G_n^S$ equals column n_{k+1-v} of G_n, for $v = 1, 2, \ldots, k$;

the other columns (if any) of $G_n G_n^S$ are zero columns. (13)

Although $G_n G_n^S$ is symmetric (for all positive integers n) for the matrices defined in Examples 1-5, $G_n G_n^S$ is not symmetric for $G_n = E_3$ given below.

$$E_3 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_3^S = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_3 E_3^S = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Theorem 2: The matrix $G_n G_n^S$ is symmetric for $n = 1, 2, 3, \ldots$ if and only if $(\alpha *_G \beta)(n) = (A *_G B)(n)$ for all arithmetic functions α and β, and for all positive integers n.

Proof:

1. Assume that $G_n G_n^S$ is symmetric for $n = 1, 2, 3, \ldots$. This and the symmetry of G_n^S imply that $(G_n G_n^S)^t = G_n (G_n^S)^t$. In view of (9), (10), (11), and (12), we have

$$\begin{aligned}(A *_G B)(n) &= A_{[n]} G_n^S \left(B_{[n]}\right)^t \\ &= \alpha_{[n]} G_n G_n^S \left(\beta_{[n]} G_n^{-1}\right)^t \\ &= \alpha_{[n]} G_n^S (G_n)^t (G_n^{-1})^t \left(\beta_{[n]}\right)^t \\ &= \alpha_{[n]} G_n^S \left(\beta_{[n]}\right)^t \\ &= (\alpha *_G \beta)(n), \quad n = 1, 2, 3, \ldots.\end{aligned}$$

2. Assume that there is a positive integer n such that $G_n G_n^S$ is not symmetric. Then $G_n G_n^S \neq (G_n G_n^S)^t$ implies that $G_n G_n^S (G_n^{-1})^t \neq G_n^S$ and that $(A *_G B)(n) = \alpha_{[n]} G_n G_n^S (G_n^{-1})^t \left(\beta_{[n]}\right)^t$ and $(\alpha *_G \beta)(n)$ are not identically equal. Therefore, there exist arithmetic functions α and β such that

$$(A *_G B)(n) \neq (\alpha *_G \beta)(n).$$

This completes the proof of the theorem. ∎

Next, we give an application of this theorem.

Example 6: Since $P_n P_n^S$ is symmetric for $n = 1, 2, 3, \ldots$ for P in Example 5, we can apply Theorem 2 with $n = 2t - 1$ (for t a positive integer), $\alpha = \nu_1$, $\beta(2k - 1) = k$ for $k = 1, 2, \ldots, t$, to obtain the identity

$$\sum_{k=1}^{t} \nu_2(k) = \sum_{k=1}^{t} (2k - 1)(t - k + 1),$$

which can be expressed in the form

$$t^3 = \sum_{k=1}^{t} \nu_2(k) + \sum_{k=1}^{t-1} k(2k + 1).$$

5. A GENERAL EULER FUNCTION

Assume that the matrix G generates the convolution $*_G$. In §3, we defined a general Moebius function μ_G and obtained a generalization of (3). In this section, we define a general Euler function ϕ_G for G such that $*_G$ satisfies (6) and (7), and derive a generalization of (4).

First, we consider the property

$$G_n G_n^S \text{ is symmetric for } n = 1, 2, 3, \ldots \tag{14}$$

and some preliminary theorems.

Theorem 3: Property (7) implies Property (14).

Proof: Assume that $G_n G_n^S = (h_{ij})$ is not symmetric.

Case 1: Suppose that column w of $G_n G_n^S$ is a zero column and that $h_{wq} = 1$ for some $q \in \{1, 2, \ldots, n\}$. By (13), $g_{wn} = 0$ and $q \in \{n_1, \ldots, n_k\}$; say $q = n_{k+1-t}$. Then $g_{wn_t} = 1 = g_{n_t n} = g_{n_t n_t}$ and $((\alpha *_G \beta) *_G \gamma)(n)$ has a term with factor $\alpha(w)$; but $(\alpha *_G (\beta *_G \gamma))(n)$ has no term with factor $\alpha(w)$ and (7) is false.

Case 2: Suppose that $h_{n_s n_r} = 0$ and $h_{n_r n_s} = 1$, where n_s and n_r belong to $\{n_1, \ldots, n_k\}$. Then $g_{n_s n_{k+1-r}} = 0$, $g_{n_r n_{k+1-s}} = 1$, and $g_{n_s n} = 1 = g_{n_r n}$. Therefore, $(\alpha *_G \beta)(n_{k+1-s})\gamma(n_s)$ has a term with factors $\alpha(n_r)$ and $\gamma(n_s)$, but $\alpha(n_r)(\beta *_G \gamma)(n_{k+1-r})$ has no term with a $\gamma(n_s)$ factor. Again, (7) is false. ■

Theorem 4: Property (14) implies Property (8).

Proof: Assume that (8) is false and let t and r be defined as in the proof of Theorem 1. Column t of $G_n G_n^S$ is a zero column (since $g_{tn} = 0$); but a 1 entry appears in row t of $G_n G_n^S$ (because $g_{tr} = 1 = g_{rn}$), so that $G_n G_n^S$ is not symmetric. ■

We note that (7) implies (8) and (14), and that (14) implies (8); there are no other implications among the properties (6), (7), (8), and (14) (as will be shown in §5).

It follows from (9) that $A = \nu_0 *_G \alpha$. If G and $*_G$ satisfy (6) and (7), then (by Theorems 3 and 2) we have $(\alpha *_G \beta)(n) = (\alpha *_G \nu_0 *_G B)(n)$ for all arithmetic functions α and β and for $n = 1, 2, 3, \ldots$. Therefore, we have

$$\beta(n) = (\nu_0 *_G B)(n);$$

and

$$B(n) = (\beta *_G \mu_G)(n) \tag{15}$$

for all arithmetic functions β and for $n = 1, 2, 3, \ldots$ follows by Moebius inversion.

Theorem 5: If properties (6) and (7) hold for G and $*_G$, then

$$\overline{g}_{n_v n} = \mu_G(n_{k+1-v}),\quad v = 1, 2, \ldots, k.$$

Proof: Define the arithmetic functions β_v, $v = 1, 2, \ldots, k$, by $\beta_v(n) = 1$ if $n = n_v$, $\beta_v(n) = 0$ otherwise. Property (15) implies that

$$\sum_{i=1}^{n} \beta(i)\overline{g}_{in} = \sum_{v=1}^{k} \beta(n_v)\mu_G(n_{k+1-v}) \tag{16}$$

for all arithmetic functions β and for $n = 1, 2, 3, \ldots$. Let $G = *_G$ in (16) to obtain $\overline{g}_{n_v n} = \mu_G(n_{k+1-v})$; this is valid for $v = 1, 2, \ldots, k$. ■

For G and $*_G$ which satisfy (6) and (7) we define the general Euler function ϕ_G by $\phi_G = \nu_1 *_G \mu_G$. We can now generalize (4).

Theorem 6: If G and $*_G$ satisfy (6) and (7), then $\nu_{1[n]} G_n^{-1} = \phi_{G[n]}$.

Proof: This is a direct consequence of Theorem 5 and Property (8) (which follow from (6), (7), and Theorems 3 and 4). ■

Other general functions such as τ_G and σ_G can be defined analogously.

6. REMARKS

First, we show that there are no implications among properties (6), (7), (8), and (14) except (7) implies (8) and (14), and (14) implies (8). If R_5 is as shown and $R = (r_{ij})$ is defined for $i > 5$ and $j > 5$ by $r_{ij} = 1$ if $i = j$ or $i = 1$, $r_{ij} = 0$ otherwise, then R satisfies (6) but not (7), (8), and (14). The matrix P defined in

$$R_5 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 0 & 0 \\ & & 1 & 1 & 0 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}, \qquad M_5 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 1 & 1 \\ & & 1 & 1 & 1 \\ & & & 1 & 1 \\ & & & & 1 \end{bmatrix}.$$

Example 5 satisfies (7), (8), and (14) but not (6). A matrix $M = (m_{ij})$ which satisfies (8) but not (7) and (14) can be defined for $i > 5$ and $j > 5$ by $m_{ij} = 1$ if $i = j$, $m_{ij} = 0$ otherwise, with M_5 as shown. If K_{10} is as shown and $K = (k_{ij})$ is defined for $i > 10$ and $j > 10$ by $k_{ij} = 1$ if $i = j$, $k_{ij} = 0$ otherwise, then (14) holds, but (7) is false since, for example,

$$((\nu_1 *_K \nu_1) *_K \nu_0)(10) \neq (\nu_1 *_K (\nu_1 *_K \nu_0))(10).$$

$$K_{10} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ & & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 & 0 & 1 & 1 \\ & & & & & 1 & 0 & 0 & 0 & 1 \\ & & & & & & 1 & 0 & 0 & 0 \\ & & & & & & & 1 & 0 & 0 \\ & & & & & & & & 1 & 1 \\ & & & & & & & & & 1 \end{bmatrix}.$$

Properties (6), (7), (8), and (14) all hold for the matrices (and generated convolutions) in Examples 1-4 as well as for those defined in our concluding example.

Example 7: Let $\hat{F} = \{1, 2, 3, 5, 8, \ldots\}$ be the set of positive Fibonacci numbers. Define $\hat{F} = (f_{ij})$ by $f_{ij} = 1$ if $i = j$ or if $i < j$ and $i \varepsilon \hat{F}$, $f_{ij} = 0$ otherwise. $\hat{F}$ can be replaced by any finite or infinite set of positive integers which includes 1, and properties (6), (7), (8), and (14) will be satisfied. If $\hat{F}$ is replaced by the set of all positive integers, we obtain the matrix C in Example 3.

REFERENCES

1. L. Carlitz, "Arithmetic Functions in an Unusual Setting," *American Math. Monthly*, Vol. 73 (1966), pp. 582-590.
2. E. E. Guerin, "Matrices and Arithmetic Functions" (submitted).
3. D. A. Smith, "Incidence Functions as Generalized Arithmetic Functions, I," *Duke Math. Journal*, Vol. 34 (1967), pp. 617-633.

*****