Möbius Inversion of Random Acyclic Directed Graphs

By Joel E. Cohen

Suppose a random acyclic digraph has adjacency matrix $\mathbf{A}$ with independent columns or independent rows. Then the mean Möbius inverse of the zeta matrix $I+\mathbf{A}$ is the Möbius inverse of the mean zeta matrix, i.e., $E[(I+\mathbf{A})^{-1}] = [I+E(\mathbf{A})]^{-1}$.

The purpose of this note is to show that, under natural conditions, the mean Möbius inverse of a random acyclic directed graph (digraph) equals the Möbius inverse of the mean acyclic digraph.

Let the vertex set V be $\{1,\dots,n\}$ for a fixed integer n, $1<n<\infty$, and let R be a subset of $V\times V$. An element $(i,j)\in R$ is called an arc from i to j. A digraph D is an ordered pair $D=(V,R)$ of vertices and arcs. A topologically ordered acyclic digraph (TOAD) is a digraph $D=(V,R)$ such that every arc (i,j) in R satisfies $i<j$. It is well known that every acyclic digraph can be converted to a TOAD by permuting the labels of the vertices, and conversely every acyclic digraph can be obtained by permuting the labels of the vertices of a TOAD. The adjacency matrix $A=A(D)=A(V,R)$ of any digraph $D=(V,R)$ is an $n\times n$ matrix such that $a_{ij}=1$ if $(i,j)\in R$, $a_{ij}=0$ if $(i,j)\notin R$. It is also well known that (V,R) is a TOAD if and only if $A(V,R)$ is strictly upper triangular, i.e., $a_{ij}=0$ whenever $i\geq j$. (See [4] for background on digraphs.)

The zeta function of any acyclic digraph D with adjacency matrix A is defined by $\zeta=I+A$, where I is the $n\times n$ identity matrix. ζ is the adjacency matrix of the digraph formed from D by adjoining loops to each vertex, i.e., by adjoining all the arcs (i,i) where $i\in V$. The Möbius inverse $\mu=\zeta^{-1}$ exists, because if P is a permutation matrix such that PAP^{-1} is strictly upper

Address for correspondence: Professor Joel E. Cohen, Rockefeller University, 1230 York Avenue, Box 20, New York, NY 10021-6399.

Published by Elsevier Science Publishing Co., Inc. 0022-2526/91/$3.50

triangular, then $\det(I+A)=\det[P(I+A)P^{-1}]=\det(I+PAP^{-1})=1$. (Recall that for a permutation matrix P, $P^{-1}=P^T$, so PAP^{-1} is the matrix obtained by relabeling the indices of the rows and columns of A according to P.) (See [5; 2, Chapter 2; 3, Chapter 25] for background on Möbius inverses.)

Let S be the set of all strictly upper triangular $0-1$ $n\times n$ matrices, and let U be the set of all matrices PAP^T, where P is a permutation matrix and $A\in S$. The matrices in S are exactly the adjacency matrices of the set of all TOADs, and the matrices in U are exactly the adjacency matrices of the set of all acyclic digraphs.

A random acyclic digraph is specified by a probability distribution on U. Specifically, if $\mathbf{A}$ denotes the random adjacency matrix of a random acyclic digraph $\mathbf{D}$, then for every $A\in U$, $p(A)=P\{\mathbf{A}=A\}$. The mean adjacency matrix of $\mathbf{D}$ is $E(\mathbf{A})=\Sigma_{A\in U}Ap(A)$. The mean Möbius inverse of $\mathbf{D}$ is $E(\boldsymbol{\mu})=\Sigma_{A\in U}(I+A)^{-1}p(A)$. Under natural conditions, stated in Theorem 2, $E(\boldsymbol{\mu})=[I+E(\mathbf{A})]^{-1}$. This follows from a slightly more general result, stated as Theorem 1.

Let $\mathbf{M}$ be a random $n\times n$ matrix (implicitly, a space of $n\times n$ matrices together with a probability measure on that space). Say that $\mathbf{M}=(\mathbf{m}_{ij})$ has independent columns if and only if, for all j,k such that $1\le j<k\le n$, the vector consisting of column j of $\mathbf{M}$ and the vector consisting of column k of $\mathbf{M}$ are independent. (Arbitrary dependence within any column is allowed.) Say that a random matrix has independent rows if its transpose has independent columns. If the rows or columns of $\mathbf{M}$ are independent, so are those of $P\mathbf{M}P^T$, for any permutation matrix P.

As usual, "a.s." means "almost surely."

For any deterministic matrix $M=(M_{ij})$, the skeleton of M is another deterministic matrix $H=(h_{ij})$ such that $h_{ij}=1$ if $m_{ij}\ne 0$, $h_{ij}=0$ if $m_{ij}=0$. For any random matrix $\mathbf{M}$, define the *movie* of $\mathbf{M}$ to be the random matrix $\mathbf{H}$ formed by taking the skeleton of each realization of $\mathbf{M}$. For any random matrix $\mathbf{M}$, define the *still* of $\mathbf{M}$ to be the deterministic matrix H defined by $h_{ij}=0$ if $m_{ij}=0$ a.s., $h_{ij}=1$ if $P\{m_{ij}\ne 0\}>0$. For example, if

$$M_1=\begin{pmatrix}0&2\\0&0\end{pmatrix},\qquad M_2=\begin{pmatrix}0&0\\3&0\end{pmatrix},$$

and $P\{\mathbf{M}=M_1\}=\frac{1}{4}$, $P\{\mathbf{M}=M_2\}=\frac{3}{4}$, then the movie

$$\mathbf{H}=\begin{pmatrix}0&1\\0&0\end{pmatrix}\text{ whenever }\mathbf{M}=M_1,\qquad \mathbf{H}=\begin{pmatrix}0&0\\1&0\end{pmatrix}\text{ whenever }\mathbf{M}=M_2,$$

and the still

$$H=\begin{pmatrix}0&1\\1&0\end{pmatrix}.$$

Though $\mathbf{H}$ is acyclic a.s., the skeleton of $E(\mathbf{H})$, namely H, is not acyclic.

Define a deterministic matrix M to be nilpotent if there exists a positive integer k such that $M^k = 0$. It is well known that M is nilpotent if and only if the skeleton of M is in U, i.e., if and only if PMP^T is strictly upper triangular for some permutation matrix P. The previous example shows that it is possible to have $\mathbf{M}$ be nilpotent a.s. while the still of $\mathbf{M}$ is not nilpotent. However, if $\mathbf{M}$ has independent rows or columns, such a possibility is excluded.

LEMMA. *Let* $\mathbf{M}$ *be a complex-valued random matrix with independent rows or independent columns and such that* $\mathbf{M}$ *is nilpotent a.s. Then the still of* $\mathbf{M}$ *is nilpotent.*

Proof: Let $H = (h_{ij})$ be the still of $\mathbf{M}$. Then $P\{\mathbf{M} \text{ is nilpotent}\} = 1$ implies that for every k, $1 \le k \le n$, and for every set $\{i_1, \ldots, i_k\}$ of k distinct elements of $V = \{1, \ldots, n\}$,

$$P\{\mathbf{m}_{i_1 i_2}\mathbf{m}_{i_2 i_3} \cdots \mathbf{m}_{i_k i_1} \neq 0\} = 0.$$

Since the rows or columns of $\mathbf{M}$ are independent,

$$0 = P\{\mathbf{m}_{i_1 i_2} \cdots \mathbf{m}_{i_k i_1} \neq 0\} = P\{\mathbf{m}_{i_1 i_2} \neq 0\} \cdots P\{\mathbf{m}_{i_k i_1} \neq 0\},$$

which implies that at least one factor on the right is 0. Therefore, at least one of $h_{i_1 i_2}, h_{i_2 i_3}, \ldots, h_{i_k i_1}$ is 0. Since this is true for every set $\{i_1, \ldots, i_k\}$ of k distinct elements of V, H is nilpotent. $\square$

Define the off-diagonal part of a random matrix $\mathbf{A}$ to be the random matrix $\mathbf{B}$ such that $\mathbf{b}_{ij} = \mathbf{a}_{ij}$ a.s. for all (i, j) with $i \neq j$, and $\mathbf{b}_{ii} = 0$ a.s. for all $i \in V$.

THEOREM 1. *Let* $\mathbf{M}$ *be a complex-valued random matrix such that*

(i) *the expectation* $E(\mathbf{M})$ *exists*;
(ii) *the off-diagonal part of* $\mathbf{M}$ *is a.s. nilpotent, i.e., the movie of the off-diagonal part of* $\mathbf{M}$ *is a.s. nilpotent*;
(iii) $\mathbf{m}_{ii} = c_i$ *a.s., where* $c_i \neq 0$ *is a nonzero constant*;
(iv) $\mathbf{M}$ *has independent columns or* $\mathbf{M}$ *has independent rows.*

Then $E(\mathbf{M}^{-1})$ *exists and* $E(\mathbf{M}^{-1}) = [E(\mathbf{M})]^{-1}$.

Proof: By the Lemma, the still of the off-diagonal part of $\mathbf{M}$ is nilpotent, and therefore so is the expectation of the off-diagonal part of $\mathbf{M}$. Hence $E(\mathbf{M})$ is nonsingular.

Let $K = (k_{ij})$ have elements $\delta_{ij} c_i$, where δ_{ij} is Kronecker's delta, $\delta_{ii} = 1$, $\delta_{ij} = 0$ if $i \neq j$. Then $\mathbf{H} = K^{-1}\mathbf{M}$ has all diagonal elements equal to 1 a.s. Let $I - \mathbf{H} = \mathbf{L}$. Then $\mathbf{L}$ has a.s. the same movie as the off-diagonal part of $\mathbf{M}$ and is a.s. nilpotent, so a.s. $\mathbf{L}^n = 0$. Therefore, $(I - \mathbf{L})^{-1} = I + \mathbf{L} + \mathbf{L}^2 + \cdots + \mathbf{L}^{n-1}$ a.s.

Hence a.s.

$$\mathbf{M}^{-1} = (K\mathbf{H})^{-1} = \mathbf{H}^{-1}K^{-1} = (I-\mathbf{L})^{-1}K^{-1}$$
$$= (I+\mathbf{L}+\mathbf{L}^2+\cdots+\mathbf{L}^{n-1})K^{-1},$$

so if $E(\mathbf{M}^{-1})$ exists, it must be

$$E(\mathbf{M}^{-1}) = \left[I+E(\mathbf{L})+E(\mathbf{L}^2)+\cdots+E(\mathbf{L}^{n-1})\right]K^{-1}.$$

Now for $k=2,\dots,n-1$,

$$(\mathbf{L}^k)_{ij} = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\cdots\sum_{i_{k-1}=1}^{n}\mathbf{L}_{i,i_1}\mathbf{L}_{i_1i_2}\cdots\mathbf{L}_{i_{k-1},j}.$$

Because $\mathbf{L}$ is a.s. nilpotent, the only terms on the right that are not a.s. 0 are those in which $i,i_1,i_2,\dots,i_{k-1},j$ are all distinct. Then, since $\mathbf{L}$ has independent columns or rows (inherited from $\mathbf{M}$ and $\mathbf{H}$),

$$E\left[(\mathbf{L}^k)_{ij}\right] = \sum_{\{i,i_1,\dots,i_{k-1},j\}\text{ all distinct}} E\left(\mathbf{L}_{i,i_1}\mathbf{L}_{i_1i_2}\cdots\mathbf{L}_{i_{k-1},j}\right)$$
$$= \sum_{\{i,i_1,\dots,i_{k-1},j\}\text{ all distinct}} E\left(\mathbf{L}_{i,i_1}\right)E\left(\mathbf{L}_{i_1i_2}\right)\cdots E\left(\mathbf{L}_{i_{k-1},j}\right);$$

hence $E(\mathbf{L}^k)=[E(\mathbf{L})]^k$ if $E(\mathbf{L})$ exists, and $E(\mathbf{L})=I-K^{-1}E(\mathbf{M})$ does exist by (i). Thus $E(\mathbf{M}^{-1})$ exists and equals

$$E(\mathbf{M}^{-1}) = \{I+E(\mathbf{L})+[E(\mathbf{L})]^2+\cdots+[E(\mathbf{L})]^{n-1}\}K^{-1}$$
$$= [I-E(\mathbf{L})]^{-1}K^{-1} = [K(I-E(\mathbf{L}))]^{-1}$$
$$= [E(\mathbf{M})]^{-1}. \qquad \square$$

By contrast with Theorem 1, if $\mathbf{X}$ is a nondegenerate random variable such that $\mathbf{X}>0$ a.s. and $E(\mathbf{X})$ and $E(\mathbf{X}^{-1})$ exists, then $[E(\mathbf{X})]^{-1}<E(\mathbf{X})^{-1}$. [Since $f(x)=x^{-1}$ is strictly convex on $(0,\infty)$, the inequality follows by Jensen's inequality.] The matrix equality obtained in Theorem 1 differs from the scalar inequality because, of course, the 1×1 case of the matrix $\mathbf{M}$ is a.s. a constant, not a nondegenerate scalar random variable; for a constant scalar or degenerate random variable $\mathbf{X}$, $[E(\mathbf{X})]^{-1}=E(\mathbf{X}^{-1})=\mathbf{X}^{-1}$ a.s.

THEOREM 2. *Suppose a random acyclic digraph has adjacency matrix* $\mathbf{A}$ *with independent columns or independent rows. Then* $E(\mathbf{A})$ *exists and the mean Möbius inverse is* $E[(I+\mathbf{A})^{-1}]=[I+E(\mathbf{A})]^{-1}$.

Proof: $E(\mathbf{A})$ exists because the elements of $\mathbf{A}$ are drawn from $\{0,1\}$, so $\boldsymbol{\zeta} = I + \mathbf{A}$ satisfies hypothesis (i) of Theorem 1. The off-diagonal part of the zeta matrix $\boldsymbol{\zeta} = I + \mathbf{A}$ is just the adjacency matrix $\mathbf{A}$, so $\boldsymbol{\zeta} = I + \mathbf{A}$ satisfies (ii), (iii) because $\boldsymbol{\zeta}_{ii} = 1$ a.s., and (iv) by assumption. The conclusion of Theorem 2 then follows from Theorem 1. □

Example of Theorem 2 (The cascade model [1]): Suppose $\mathbf{a}_{ij} = 0$ a.s. if $i \geq j$, while $\mathbf{a}_{ij} = 1$ with probability p and $\mathbf{a}_{ij} = 0$ with probability $q = 1 - p$, independently for all (i,j) with $i < j$, where $0 < p < 1$. Then $E(\mathbf{a}_{ij}) = pJ_{\{i<j\}}$, where $J_{\{i<j\}} = 1$ if $i < j$, $J_{\{i<j\}} = 0$ if $i \geq j$. Let $M = (m_{ij}) = E(\boldsymbol{\mu}) = E[(I + \mathbf{A})^{-1}] = [I + E(\mathbf{A})]^{-1}$. Then it is easy to check that

$$m_{ij} = \begin{cases} 0 & \text{if } i > j, \\ 1 & \text{if } i = j, \\ -p(1-p)^{j-i-1} & \text{if } i < j. \end{cases}$$

For example, if $n = 4$, then

$$E(\mathbf{A}) = \begin{pmatrix} 0 & p & p & p \\ 0 & 0 & p & p \\ 0 & 0 & 0 & p \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$E(\boldsymbol{\mu}) = [I + E(\mathbf{A})]^{-1} = \begin{pmatrix} 1 & -p & -p(1-p) & -p(1-p)^2 \\ 0 & 1 & -p & -p(1-p) \\ 0 & 0 & 1 & -p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

In the ecological interpretation of the cascade model [1] the elements of V represent groups of organisms called trophic species, $\mathbf{a}_{ij} = 1$ means species j eats species i, and $\mathbf{a}_{ij} = 0$ means species j does not eat species i. The TOAD specified by $\mathbf{A}$ is called a food web. Let $x^T = (x_1, \ldots, x_n)$ and $y^T = (y_1, \ldots, y_n)$ be row vectors such that $y^T = x^T(I + \mathbf{A})$, i.e., y_j is the sum of x_j plus all the x_i such that j eats i according to $\mathbf{A}$. Then $y^T(I + \mathbf{A})^{-1} = x^T$. Now suppose y^T is fixed, e.g., y^T can be measured directly with negligible error. Then $E(x^T) = y^T E[(I + \mathbf{A})^{-1}] = y^T M$ provides a way of estimating the mean of x^T from measurements of y^T and the average structure of the food web; the latter may be derived from the cascade model in the absence of more detailed data.

Acknowledgments

I am grateful for the partial support of NSF grant BSR 87–05047, the hospitality of Mr. and Mrs. William T. Golden, and the hospitality of the Department of

Entomology, International Rice Research Institute, The Philippines (K. L. Heong, Acting Head).

References

1. J. E. Cohen and C. M. Newman, A stochastic theory of community food webs: I. Models and aggregated data, *Proc. Roy. Soc. London Ser. B* 224:421–448 (1985).
2. M. Hall, *Combinatorial Theory*, 2nd ed., Wiley, New York, 1986.
3. A. Nijenhuis and H. S. Wilf, *Combinatorial Algorithms for Computers and Calculators*, 2nd ed., Academic, New York, 1978.
4. D. F. Robinson and L. R. Foulds, *Digraphs: Theory and Techniques.* Gordon and Breach, London, 1980.
5. G. C. Rota, On the foundations of combinatorial theory. I. Theory of Möbius functions, *Z. Wahrsch. Verw. Gebiete* 2:340–368 (1964).

The Rockefeller University

(Received March 22, 1990)