ISSN: 0025-570X (Print) 1930-0980 (Online) Journal homepage: https://maa.tandfonline.com/loi/umma20

Applications of Graph Theory in Linear Algebra

Michael Doob

Applications of Graph Theory in Linear Algebra

Graph-theoretic methods can be used to prove theorems in linear algebra.

MICHAEL DOOB
The University of Manitoba
Winnipeg, Manitoba, Canada R3T 2N2

Graph theory has existed for many years not only as an area of mathematical study but also as an intuitive and illustrative tool. The use of graphs in wiring diagrams is a straightforward representation of the physical elements of an electrical circuit; a street map is also a graph with the streets as edges, intersections of streets as vertices, and street names as labels of the edges. The graphs resemble the physical object that they represent in these cases, and so the application (and sometimes the genesis) of the graph-theoretic ideas is immediate. A flow diagram of a computer program and a road map with one way streets are examples of graphs which contain the concept of direction or flow to the edges; these are called directed graphs.

There are applications of graphs and directed graphs in almost all areas of the physical sciences and mathematics, many of them known for fifty years or more, but very few of these ideas have percolated down to the undergraduate student. The purpose of this article is to apply graph-theoretic ideas to some of the fundamental topics in linear algebra. While there are many such applications, we shall focus on only two of the most elementary ones, i.e., matrix multiplication and the theory of determinants. The presentation is usable as a supplement to the usual classroom lectures; in fact, this paper grew out of notes used in a linear algebra course for sophomores.

The main tool to be used is the directed graph. Intuitively this can be thought of as a set of points (or vertices) with arrows (or arcs) joining some of the points. A label may be put on an arc. More formally, a **digraph** consists of a set of vertices V and a subset of ordered pairs of vertices called the **arcs**. A **labelling** of the digraph is a function from the arcs to the real numbers. A labelled digraph is usually visualized by considering the vertices as points with arcs as arrows going from vertex i to vertex j whenever (i, j) belongs to the sets of arcs. The ith vertex of the arc (i, j) is called its **initial vertex**, while the jth vertex is called its **terminal vertex**. The arc is then given a label which is the image of that arc under the labelling function. When the initial and terminal vertices are identical, the arc is called a **loop**. We shall sometimes say that an arc "goes out of" its initial vertex and that it "goes into" its terminal vertex. The number of arcs that go out of a vertex is called its **outdegree**, and the number of arcs that go into that vertex is called its **indegree**.

A digraph G is a **subgraph** of a digraph H if the vertices and arcs of G are contained in the set of vertices and set of arcs of H. One type of subgraph of H is a **walk**: this consists of a sequence of vertices $v_0, v_1, v_2, \ldots, v_n$ such that (v_{i-1}, v_i) is an arc in H for $i = 1, 2, \ldots, n$. We use $(v_0, v_1, v_2, \ldots, v_n)$ to denote such a walk (which is consistent with the notation for an arc); we then say that the **length** of the walk is n. A **path** is a walk where all the vertices are distinct, and a **cycle** has $v_1, v_2, \ldots, v_n$ distinct and $v_0 = v_n$. Another type of subgraph is a **factor**; it has the indegree and outdegree of each vertex equal to one. A moment's thought reveals that a factor is simply a (vertex) disjoint union of cycles. The **weight** of a subgraph, $W(G)$, is the product of the labels of the arcs in that subgraph. Weights of walks, paths, cycles, and factors are similarly defined.

Matrix multiplication and the König digraph

In this section we examine a particular directed graph associated with an $m \times n$ matrix A. The multiplication of two matrices is equivalent to "glueing" their digraphs together to form a single graph. An entry in the product matrix is then related to the weights of certain paths in the new graph. Most standard proofs about matrix multiplication involve the manipulation of subscripts and/or the interchanging of summations. These techniques, while valid, tend to obscure the underlying ideas; this is avoided by the graph-theoretic approach.

Let $A = (a_{ij})$ be a matrix with m rows and n columns. The **König digraph**, $G(A)$, is a labelled digraph with $m + n$ vertices—m of these vertices correspond to the rows and n correspond to the columns. The arc from vertex i to vertex j has a_{ij} as its label. When drawing the graph we shall omit arcs with a zero label; this will clarify the relationships between several matrix and graphic operations. As a further convention, we shall draw the row vertices on the left and the column vertices on the right. The top vertex will correspond to the first row or column, the one below it to the second, etc. An illustration of a matrix A and its König digraph $G(A)$ is shown in FIGURE 1. D. König [5] used this digraph in his well-known and fundamental book in 1936; he referred to it tangentially even earlier [4] in 1916.

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & -1 \end{pmatrix}$$

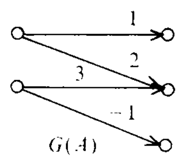

FIGURE 1. A 2×3 matrix A and its König digraph $G(A)$.

If A and B are both $m \times n$ matrices, then $G(A)$ and $G(B)$ are essentially the same digraph with different labels. The matrix sum $A + B$ is defined, and the digraph $G(A + B)$ obviously is obtained by adding the weights on the corresponding arcs of $G(A)$ and $G(B)$. The usual additive properties of matrices carry over directly to the digraphs. The proofs of these familiar properties, e.g., the associative, commutative, and distributive laws, are identical to the usual ones, but rather than focusing on a particular row and column entry in a matrix, a label of a particular arc is used to determine the validity of these properties. The multiplication of a matrix by a scalar is also easy to visualize. The digraph $G(rA)$ is obtained from $G(A)$ by multiplying every arc label by r. The usual properties of scalar multiplication are also easily verified.

Given an $m \times n$ matrix A and an $n \times r$ matrix B, the **concatenation graph** $G(A) * G(B)$ is defined by taking the digraphs $G(A)$ and $G(B)$ and identifying the n column vertices of $G(A)$ with the n row vertices of $G(B)$. An example of the concatenation of two graphs is given in FIGURE 2.

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 2 & 0 \\ 4 & -1 \\ 0 & -2 \end{pmatrix}$$

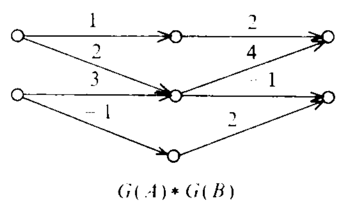

FIGURE 2. The concatenation of two König digraphs.

THEOREM 1. *The (i, j) entry of the matrix product AB is equal to the sum of the weights of the paths in $G(A) * G(B)$ from the ith row vertex of $G(A)$ to the jth column vertex of $G(B)$.*

Proof. Each path of length two from vertex i to vertex j passes through a unique vertex k. By definition, the weight of the path is the product of the labels of its arcs, which is $a_{ik}b_{kj}$. Summing over all possible k gives the desired result.

Note that the exclusion of edges of weight zero from the digraph according to our convention still leaves the theorem valid. We see, for example, that for the digraph in FIGURE 2, the (1,2) entry of the product AB is -2 since there is only one path between the appropriate vertices.

Several properties of matrix multiplication are immediate consequences of Theorem 1. To show *matrix multiplication is associative*, assume that matrices A, B, and C are of appropriate orders, and consider the graph $G(A) * G(B) * G(C)$. Then the (i, j) entry of $(AB)C$ and $A(BC)$ are clearly both equal to the sum of the weights of the paths of length three from the ith row vertex of $G(A)$ to the jth column vertex of $G(C)$. It is an easy exercise to show that *matrix multiplication distributes over addition* by looking at a particular path (i, j, k) in the König digraphs of $G(A) * G(B)$, $G(B) * G(C)$, and $G(A) * G(B+C)$.

If P is a permutation matrix, then $G(P)$ consists of a set of arcs with no common vertices. In other words, the outdegree of each row vertex and the indegree of each column vertex is equal to one, and all labels of arcs are 1. *The product of permutation matrices is a permutation matrix.* Look at the concatenation digraph! Likewise, the digraph of a diagonal matrix makes the following facts trivial. *The product of diagonal matrices is a diagonal matrix. Each diagonal entry of the product is the product of the corresponding diagonal entries in the factors.*

A matrix is upper (lower) triangular if $a_{ij} = 0$ whenever $i > j$ $(i < j)$. Thus $G(A)$ is the König digraph of an upper (lower) triangular matrix if and only if all the arcs in the drawing are horizontal or downward (upward). *The product of upper (lower) triangular matrices is upper (lower) triangular.* The digraph makes it trivial!

The König digraph of the transpose of A is obtained from $G(A)$ by reversing the direction of all of the arcs. If we wish to adhere to the convention of having row vertices on the left and column vertices on the right, we must then reflect the new graph with respect to a line through the column vertices (see FIGURE 3). The properties of the reversals and reflection make the following Theorem obvious.

THEOREM 2. *Let A^T be the transpose of the matrix A; then*

$$(1)\quad (A^T)^T = A,$$

$$(2)\quad (A+B)^T = A^T + B^T,$$

$$(3)\quad (cA)^T = cA^T, \text{ and}$$

$$(4)\quad (AB)^T = B^T A^T.$$

Notice how the reversal of the arcs makes the reversal of the order of the matrices A and B in Equation (4) intuitively clear.

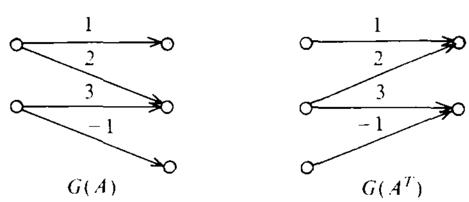

FIGURE 3

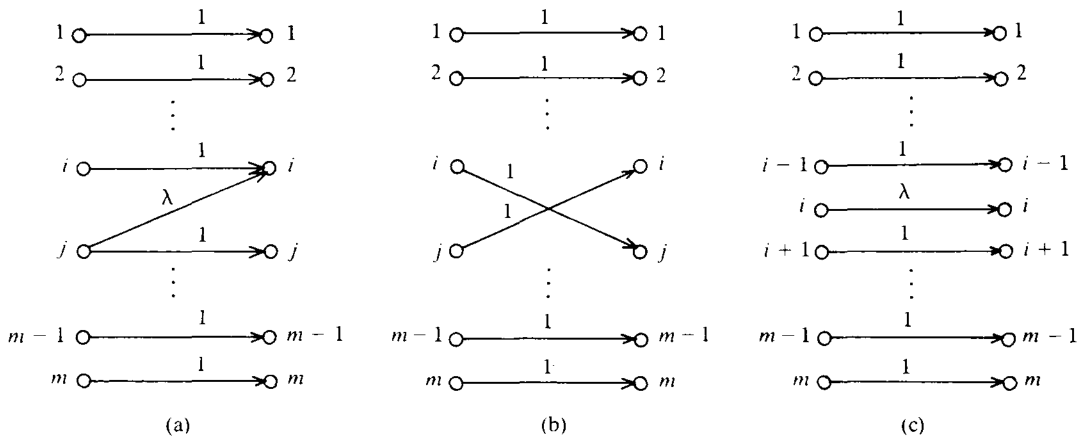

FIGURE 4. **König digraphs of the elementary row operations on an $m \times n$ matrix: (a) adds a multiple of the ith row to the jth row; (b) interchanges the ith and jth rows; (c) multiplies the ith row by λ.**

The (i, j) element of the product matrix AB is completely determined by the subgraph of $G(A) * G(B)$ consisting of the ith row vertex of $G(A)$, the jth column vertex of $G(B)$, and all paths of length two joining them. A partition of the column vertices of $G(A)$ (= the row vertices of $G(B)$) produces a partition of these paths. The validity of block multiplication is now easy to understand since each block arises from a partition of the rows and columns of the matrix.

Properties of matrices associated with elementary row operations are easily seen using König digraphs, since multiplying by a matrix on the left is nothing more than concatenation on the left by a König digraph. For example, if we wish to multiply the ith row of the $m \times n$ matrix A by λ and add it to the jth row, we simply concatenate the digraph shown in FIGURE 4(a) with $G(A)$. It is an easy exercise to construct the Konig digraphs that represent the interchange of two rows of a matrix or the multiplication of a row by λ. See FIGURE 4(b), (c). The digraphs in FIGURE 4 are called the König digraphs of the elementary row operations. A glance at these quickly shows that the König digraphs of an elementary row operation and its inverse are identical except for at most one label; in one case λ is replaced by $-\lambda$ and in the other it is replaced by λ^{-1}.

The Coates graph and the determinant

In this section we look at a different digraph associated with a square matrix. The determinant of the matrix can be computed from the weights of the cycles in this graph. This allows a relatively straightforward computation of the determinant of a matrix of arbitrary order, especially if there are many zero entries. Usually determinants are defined by making an excursion into the theory of permutations, a subject which by its nature is deeper than the determinant concept itself and necessitates a relatively difficult digression. An alternative approach is to define the determinant inductively using the cofactors, but like many inductive approaches, the results of the proofs are often believed but not really understood. The graph-theoretic approach avoids both of these pitfalls. The proof, for example, that the determinant of a matrix A and the determinant of A^T are equal tends to become lost in notation when using either the permutation or inductive approach. With the graph-theoretic approach this result is proven in one sentence!

Recall that a factor F of a digraph H is a subgraph containing all the vertices of H in which each vertex has both indegree and outdegree equal to one. In other words, it consists of a collection of disjoint cycles that go through each vertex of H. The number of cycles in the factor F is denoted $n(F)$. If the digraph H is labelled, then $W(F)$ denotes the weight of the factor.

Given a square matrix $A = (a_{ij})$ of order n, the **Coates digraph** $D(A)$ is a labelled digraph with n vertices where the arc from vertex i to vertex j has a_{ij} as a label. In FIGURE 5, the Coates

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

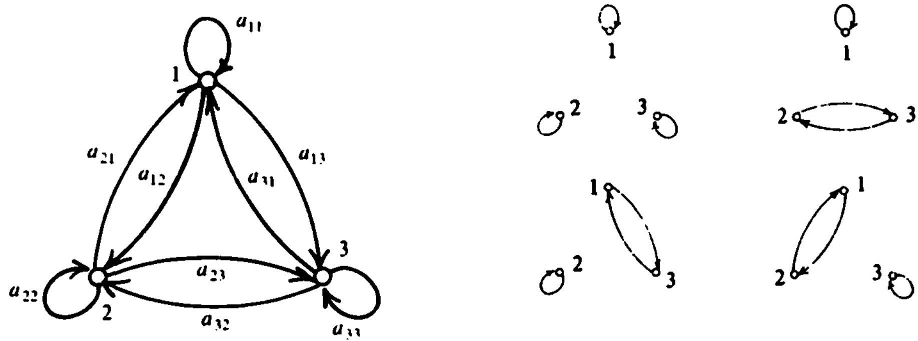

FIGURE 5. **The Coates digraph $D(A)$ of a matrix A of order $n = 3$, and the six factors of $D(A)$.**

digraph $D(A)$ of a 3×3 matrix A is shown. The digraph $D(A)$ for $n = 3$ has six factors; the two having no loops are the cycles $(1,2,3,1)$ and $(1,3,2,1)$, the three with one loop are $(1,1)$ $(2,3,2)$, $(2,2)$ $(1,3,1)$ and $(3,3)$ $(1,2,1)$; finally, $(1,1)$ $(2,2)$ $(3,3)$ consists of three loops.

Given a square matrix A of order n, $D(A)$ its Coates digraph, and $\mathbf{F}$ the set of all factors of $D(A)$, then the determinant of A is defined

$$\det A = (-1)^n \sum_{F \in \mathbf{F}} (-1)^{n(F)} W(F). \tag{1}$$

This formulation of the determinant first appeared in 1959 [**1**], although some of the relationships between graphs and determinants were known much earlier [4]. Except for the work of Cvetković [**2**], it has received attention only in a few scattered research papers. In the example of the 3×3 matrix (FIGURE 5), the factors having no loops contribute $-a_{12}a_{23}a_{31}$ and $-a_{13}a_{32}a_{21}$ to the sum, those having one loop contribute $a_{11}a_{23}a_{32}$, $a_{13}a_{22}a_{31}$, and $a_{12}a_{21}a_{33}$ to the sum, and the final factor adds $-a_{11}a_{22}a_{33}$. Since $(-1)^3 = -1$, we observe that equation (1) yields the standard result. It is straightforward to verify the validity of formula (1) for the determinant for $n = 1$, $n = 2$, and $n = 4$. For $n = 4$, we note that the set of all factors can be partitioned into 9 factors with no loops, 8 factors with one loop, 6 factors with two loops, and one factor with 4 loops.

We now prove that formulation (1) of the determinant is identical with the usual permutation definition.

THEOREM 3. *The graphic and permutation definitions of* $\det A$ *are equivalent, i.e., if* $\det A$ *is defined by equation* (1), *then*

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)},$$

where S_n is the set of all permutations of the integers $\{1, 2, \ldots, n\}$.

Proof. For any permutation σ, the set of arcs $\{(i, \sigma(i)) | i = 1, \ldots, n\}$ forms a factor F. Conversely, each factor F also corresponds to a permutation σ. Hence to complete the proof we need only show that the two definitions produce values with the same sign, i.e., that $(-1)^n(-1)^{n(F)} = \operatorname{sgn}(\sigma)$. The vertices in a cycle of the factor F correspond precisely to the integers in a cycle in the decomposition of the corresponding permutation σ. In addition, a cyclic permutation is even if and only if the corresponding cycle in the digraph has an odd number of vertices. Let $e(\sigma)$ and $o(\sigma)$ denote the number of even and odd cycles in the permutation σ (or, equivalently, the odd and even cycles in the factor). Then

$$\operatorname{sgn}(\sigma) = (-1)^{o(\sigma)}, n(F) = e(\sigma) + o(\sigma),$$

and n and $e(\sigma)$ have the same parity. Hence

$$\operatorname{sgn}(\sigma) = (-1)^{o(\sigma)} = (-1)^{n(F)-e(\sigma)} = (-1)^{e(\sigma)}(-1)^{n(F)} = (-1)^n(-1)^{n(F)}.$$

Using definition (1), here is the proof that *the determinant of a matrix and the determinant of its transpose are equal*. The Coates digraph $D(A^T)$ is obtained from $D(A)$ by reversing the orientation of each arc; this leaves the factors, the weight of each factor, and the number of cycles in each factor unchanged, and hence the determinant is unchanged.

In the last section, we considered matrices that correspond to the elementary row operations. The matrix that interchanges two rows will have Coates digraph of a single factor, i.e., $n-2$ loops and one cycle of length two. Since all weights are equal to one, the determinant is $(-1)^n(-1)^{n-1} = -1$. Adding a multiple of one row to another also corresponds to a Coates digraph with only one factor, and the determinant is 1 in that case. Finally, multiplying a row by λ corresponds to a diagonal matrix, and this matrix has a Coates digraph with only loops, hence its determinant is λ.

Consider an upper triangular matrix as a further example. As can be seen in FIGURE 6, if the vertices of its Coates digraph are placed horizontally in increasing order, then all arcs are loops or go from left to right, and hence the only possible cycles are loops. But this means that the only (nonzero) factor is (1,1) (2,2) (3,3)...(n, n), hence

$$\det A = (-1)^n(-1)^n a_{11}a_{22} \cdots a_{nn}.$$

We have proved *the determinant of an upper triangular matrix is the product of its diagonal elements*.

A slightly harder combinatorial result can be obtained if we let $C(n, r)$ denote the number of combinations of n things taken r at a time, and let A be a cyclic tridiagonal matrix, i.e., the nonzero entries of A satisfy $a_{ii} = d$, $a_{i\,i+1} = c$, or $a_{i+1\,i} = b$. The reader can prove that

$$\det A = \sum_{r=0}^{[n/2]} C(n-r, r) b^r c^r d^{n-2r}.$$

Now let us look at the effect of elementary row operations on the Coates digraph and consequently on the determinant.

THEOREM 4. *Let A be a square matrix of order n, $1 \leqslant i, j \leqslant n$, and let*
B be obtained from A by multiplying the ith row by λ,
C be obtained from A by interchanging the ith and jth rows,
E be obtained from A by adding the jth row to the ith row.

Then $\det B = \lambda \det A$,
$\det C = -\det A$, *and*
$\det E = \det A$.

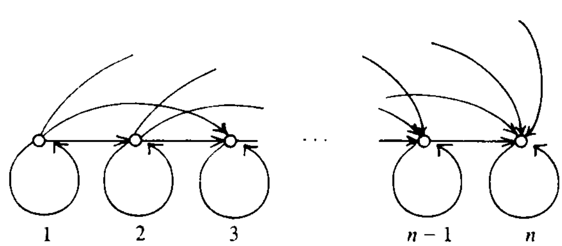

FIGURE 6. **The Coates graph of an upper triangular matrix.**

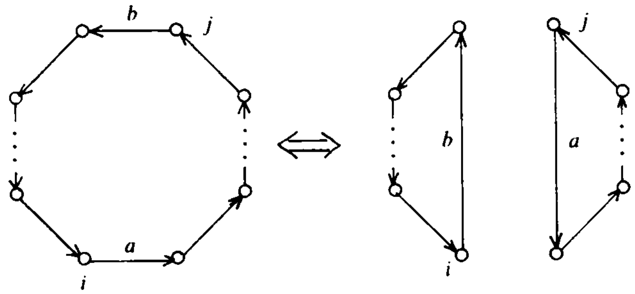

FIGURE 7. The change in a factor under row interchange.

Proof. $D(B)$ is obtained from $D(A)$ by multiplying the label of each arc going out of the ith vertex by λ. Thus F is a factor of $D(A)$ if and only if F is a factor of $D(B)$. Since the outdegree of each vertex in F is one, the weights of all arcs of F in $D(A)$ but one are the same as the weights of F in $D(B)$, the exception being that single arc going out of the ith vertex which has been multiplied by λ. Thus the weight of F in $D(B)$ is λ times that in $D(A)$, and, summing over all of the factors, $\det B = \lambda \det A$.

The Coates graphs $D(A)$ and $D(C)$ can be related as follows: we wish to take the digraph $D(A)$ and move some of the arcs. Since $a_{ik} = c_{jk}$ and $a_{jk} = c_{ik}$, $D(C)$ is obtained from $D(A)$ by taking each arc going out of the ith vertex and moving it so that it goes out of the jth vertex, and conversely, moving the arcs going out of the jth vertex so that they can then go out of the ith vertex (the terminal end of the arc remains fixed). The net change of these movements of arcs is to interchange a_{ik} and a_{jk} for every k, which, of course, is just what is desired. Now consider a particular factor F, and suppose that the labels of the arcs going out of vertex i and vertex j are a and b. How do these row interchanges affect F? It remains a factor with $W(F)$ unchanged! As can be seen in FIGURE 7, if both the ith and the jth vertices are in the same cycle to begin with, then the movement of the arcs splits the cycle into two new ones with the same weight (left to right in FIGURE 7). On the other hand, if the ith and jth vertices are in different cycles, then the movement of the arcs causes them to combine into one cycle with the same weight (right to left in FIGURE 7). In either case, $W(F)$ is unchanged and $n(F)$ is increased or decreased by one. Thus the sign of each summand changes, and $\det C = -\det A$.

Notice that if A has two identical rows, then interchanging them leaves A unchanged but reverses the sign of $\det A$, and so $\det A = 0$. Now let A' be obtained from A by replacing the ith row of A by the jth row of A (so that $\det A' = 0$). Consider a factor F of $D(E)$; it is also a factor of $D(A)$ and $D(A')$, and all the weights of the arcs are identical except for the arc in F going out of the ith vertex. The weight of this arc is $a_{ik} + a_{jk}$ in E, is a_{ik} in A, and is a_{jk} in A'. Thus the weight of F in $D(E)$ is the sum of the weight of F in $D(A)$ and the weight of F in $D(A')$. Summing over all factors F, we get $\det E = \det A + \det A' = \det A$.

If one already knows how to reduce a matrix to reduced row echelon form, then it is easy to see from Theorem 4 that $\det EA = \det E \cdot \det A$ for an elementary matrix E and consequently that $\det AB = \det A \cdot \det B$ holds in general. In the next section we shall prove this result with graph-theoretic tools.

Sometimes the determinant is defined abstractly as a function from the set of square matrices of order n to the real numbers which, viewed as an n-variable function of the columns, is alternating, multilinear, and takes the identity matrix to 1. Since $\det A^T = \det A$, we can use rows instead of columns, of course. Suppose F is a factor of a directed graph with n vertices. Then, by the argument in Theorem 4, the mapping that takes a square matrix M into $(-1)^n W(F)$ is a linear function in each row of M. Hence, summing over all rows, we see from (1) that the determinant is indeed an alternating multilinear form on the rows of M, and that the determinant of the identity matrix is 1. Since it is easily seen that such a form is unique (see [6] p. 191, for example), we now see that the graphic definition is equivalent to the abstract definition.

We now consider expansion of the determinant by cofactors. Given a square matrix A, let D_{ij} be defined as the determinant of the matrix obtained by deleting the ith row and jth column from A. The **(i,j) cofactor of A**, denoted A_{ij}, is then defined by

$$A_{ij} = (-1)^{i+j} D_{ij}.$$

Now consider the set $\mathbf{F}$ of all factors of $D(A)$ and a particular vertex i. Each factor contains a unique arc going out of the ith vertex. Let $\mathbf{F}_j$ be the set of factors containing the arc (i, j). Then clearly $\mathbf{F}$ is partitioned by $\mathbf{F}_j, j = 1,2,\dots,n$. The (i, j) cofactor can now be expressed in terms of $\mathbf{F}_j$.

THEOREM 5. *Let $A = (a_{ij})$ be a square matrix of order n, $1 \leqslant j \leqslant n$, and let $\mathbf{F}_j$ be the set of factors of $D(A)$ containing the arc (i, j). Then*

$$a_{ij}A_{ij} = (-1)^n \sum_{F \in \mathbf{F}_j} (-1)^{n(F)} W(F). \tag{2}$$

Proof. First, suppose that $i = j$, and let A' be the square matrix of order $n-1$ obtained by deleting the ith row and column from A. By definition, any F in $\mathbf{F}_i$ will contain the loop (i, i). Thus the factors in $\mathbf{F}_i$ are precisely the factors of $D(A')$ plus the loop (i, i). Since the weight of a factor in $\mathbf{F}_i$ is the product of a_{ii} and the weight of a factor in $D(A')$, the result is clear except, possibly, for the sign. The factor in $\mathbf{F}_i$ has one more cycle than the corresponding factor in $D(A')$, so an extra multiple of -1 is introduced inside the summation. However the order of $D(A)$ is one more than that of $D(A')$, and so one fewer multiple of -1 appears outside the summation; thus the theorem is valid when $i = j$.

To complete the proof of the theorem, we must see the effect of interchanging two adjacent columns of A on the left and right side of the equation (2). Let A' be obtained from A by interchanging columns r and $r+1$. We certainly have $a_{ir} = a'_{i\,r+1}$ and

$$A_{ir} = (-1)^{i+r} D_{ir} = -\left((-1)^{i+r+1} D'_{i\,r+1}\right).$$

Thus interchanging two adjacent columns causes the value of the left side of the equation (2) to change only in sign. For the right side, we proceed as in the proof of Theorem 4 and rearrange the arcs of $D(A)$ to form $D(A')$. Since we are interchanging columns, we move the terminal end rather than the initial end of the arcs. Thus each arc terminating at the rth vertex is moved so it terminates at the $r+1$st vertex and vice-versa. In a manner analogous to that illustrated in FIGURE 7, each factor F in $\mathbf{F}_r$ becomes a factor F' in $\mathbf{F}_{r+1}$ with $W(F) = W(F')$ and $|n(F) - n(F')| = 1$. Summing over all F in $\mathbf{F}_r$, we get

$$(-1)^n \sum_{F \in \mathbf{F}_r} (-1)^{n(F)} W(F) = -\left((-1)^n \sum_{F' \in \mathbf{F}_{r+1}} (-1)^{n(F')} W(F')\right).$$

Thus the right side of equation (2) also changes sign when two adjacent columns are interchanged. Hence equation (2) is valid for A if and only if it is valid for A'. In other words, if the theorem is valid for a particular i and j, then it is also valid for i and $j+1$.

Now suppose $i > j$; we then successively interchange columns j and $j+1$, $j+1$ and $j+2$, $j+2$ and $j+3$, etc., until we have finally interchanged columns $i-1$ and i. By the preceding argument, all of the resulting matrices will simultaneously satisfy or not satisfy equation (2) of the theorem. Since the final matrix is just the case where $i = j$, we see that the theorem is valid for all of the matrices, and, in particular, for A. The argument for $i < j$ is symmetric.

COROLLARY (*Expansion by cofactors*). *Let A be a square matrix of order n, $1 \leqslant i, j \leqslant n$; then*

$$\det A = \sum_{k=1}^{n} a_{ik}A_{ik} = \sum_{k=1}^{n} a_{kj}A_{kj}.$$

Proof. As was noted previously, $\mathbf{F}$ is partitioned by $\mathbf{F}_k$, $k = 1,2,\dots,n$. Thus

$$\det A = (-1)^n \sum_{F \in \mathbf{F}} (-1)^{n(F)} W(F) = \sum_{k=1}^{n} (-1)^n \sum_{F \in \mathbf{F}_k} (-1)^{n(F)} W(F) = \sum_{k=1}^{n} a_{ik} A_{ik}.$$

The Coates digraph has application in other areas of linear algebra including powers of matrices and spectral theory, although these applications generally focus on the paths in the digraph rather than on the cycles. For the sake of brevity, these applications must be omitted, but even the results presented here establish the natural connection between the Coates digraph and linear algebra.

Another view of the determinant

We now wish to return to the determinant of the product of two matrices. Our purpose is to give a graph-theoretic proof, not only as an end in itself, but also as a means of revealing some of the underlying combinatorial structure. For a square matrix of order n, we must first observe how a factor in the Coates digraph translates to the König digraph. Such a factor is simply a vertex-disjoint set of n arcs, and so, disregarding the signature for the moment, the summands of the determinant (Equation (1)) are precisely the weights of the subgraphs consisting of n vertex-disjoint arcs.

Now consider the concatenation of the König digraphs of two square matrices A and B of order n. How does this relate to the product of $\det A$ and $\det B$? If $\mathbf{F}_1$ is the set of factors of $G(A)$ and $\mathbf{F}_2$ is the set of factors of $G(B)$, then

$$\begin{aligned}\det A \cdot \det B &= \Big(\sum_{F_1 \in \mathbf{F}_1} (-1)^{n_1} W(F_1) \Big) \Big(\sum_{F_2 \in \mathbf{F}_2} (-1)^{n_2} W(F_2) \Big) \\ &= \sum_{\substack{F_1 \in \mathbf{F}_1 \\ F_2 \in \mathbf{F}_2}} (-1)^{n_1 + n_2} W(F_1) W(F_2).\end{aligned}$$

In other words, each summand corresponds to the weight of a subgraph in $G(A) * G(B)$ consisting of n vertex-disjoint paths of length 2. Disregarding the sign for the moment, we see that these subgraphs in fact correspond precisely to the summands of $\det A \cdot \det B$. What about the summands of $\det AB$? Each arc in AB is obtained from the paths of length 2 in $G(A) * G(B)$. Further, each arc of length 2 in $G(A) * G(B)$ will contribute to some element in AB and hence to $\det AB$. Again, expanding the products of sums we see that the summands of $\det AB$ are the weights of subgraphs of $G(A) * G(B)$ consisting of n paths of length 2 where the initial and terminal (but not necessarily the middle) vertices are distinct. The non-distinctness of the middle vertex accounts for the larger number of terms ($n!n^n$) in $\det AB$ as compared with the number $(n!)^2$ in $\det A \cdot \det B$.

To complete the proof we return to the postponed consideration of the signs of the terms. We shall show that each of the summands in $\det A \cdot \det B$ also appears in $\det AB$ with the same sign and that the remaining terms of $\det AB$ sum to zero. To do this we must first see a new relationship between the parity of a permutation σ and the König digraph. Let the König digraph of a permutation σ be the digraph of the corresponding permutation matrix, so the König digraph of σ simply consists of the set of arcs $(i, \sigma(i))$.

THEOREM 6. *The parity of a permutation σ and the parity of the number of arc intersections in the König digraph are the same.*

Proof. Suppose $\sigma(i) = i$ for $i = 1, 2, \ldots, n$. Then the number of intersections in the König digraph of σ is 0 and σ has even parity. If σ is an arbitrary permutation, we may proceed by uncrossing the arcs one pair at a time, while observing that the parity of both the permutation and the number of arc intersections change with each uncrossing.

Indeed, if the arcs $(i, \sigma(i))$ and $(j, \sigma(j))$ intersect, we may say without loss of generality that $i < j$ and $\sigma(i) > \sigma(j)$. Define σ' by $\sigma'(i) = \sigma(j)$, $\sigma'(j) = \sigma(i)$, and $\sigma'(k) = \sigma(k)$ for all other k. How does the number of arc intersections of σ' compare with that of σ? Any k not equal to i or j

must satisfy $1 \leqslant k < i$ or $i < k < j$, or $j < k \leqslant n$. Similarly, $1 \leqslant \sigma(k) < \sigma(j)$, $\sigma(j) < \sigma(k) < \sigma(i)$, or $\sigma(i) < \sigma(k) \leqslant n$. Thus there are nine possible configurations for k and $\sigma(k)$. For eight of them, the number of arc intersections is unchanged by uncrossing the arcs $(i, \sigma(i))$ and $(j, \sigma(j))$. In the configuration where $i < k < j$ and $\sigma(j) < \sigma(k) < \sigma(i)$ the number of intersections drops by two. Thus the total number of arc intersections has been decreased by twice the number of arcs of the form $\{(k, \sigma(k)) | i < k < j, \sigma(j) < \sigma(k) < \sigma(i)\}$ plus one, the extra intersection coming from the arcs $(i, \sigma(i))$ and $(j, \sigma(j))$ themselves. Hence the transposition (ij) that takes σ to σ' also causes a change in the parity of the number of arc intersections. Eventually, we get to the identity permutation, and hence the result is established.

COROLLARY. *The permutation σ and the number of pairs in the set $\{(i, j) | 1 \leqslant i < j \leqslant n, 1 \leqslant \sigma(j) < \sigma(i) \leqslant n\}$ have the same parity.*

The set defined in the Corollary is called the set of **inversions** of the permutation σ.

Note that in the proof of Theorem 6, each uncrossing represents a transposition, and the sequence of uncrossings that are used to obtain the identity permutation yields the product of transpositions that equals the original permutation. Thus the proof also shows that *any permutation is the product of transpositions*, and the graph yields an intuitive representation of σ as the product of transpositions.

Now let us apply these results to the products of determinants. One summand of $\det A \cdot \det B$ corresponds to n vertex-disjoint arcs of length two. The sign of this term is the product of the signs of the terms in $\det A$ and in $\det B$, each of which can be obtained from the number of arc intersections in its particular case. How does this compare with the same term as it appears in $\det AB$? The sign of this term comes from the parity of the number of intersections of the paths of length two. A pair of paths will intersect if the first arcs of the paths intersect in $G(A)$ or if the second arcs intersect in $G(B)$. But notice that if both the first arcs and second arcs intersect, then they yield a pair of non-intersecting arcs as far as $\det AB$ is concerned. Thus the number of intersecting arcs in $\det AB$ and the product of those in $\det A$ and $\det B$ have the same parity, which is what was desired.

Finally, let us take care of those extra terms that were in $\det AB$ but did not appear in $\det A \cdot \det B$. Let (i, k, l) and (j, k, m) be two paths of length two in $G(A) * G(B)$. How can these paths appear in $\det AB$? There are two ways: they come from terms of the form $(a_{ik} b_{kl})(a_{jk} b_{km})x$ or terms of the form $(a_{ik} b_{km})(a_{jk} b_{kl})x$ where x represents the remaining terms in each of the products. Hence the permutations that give rise to these paths can be paired so that they differ only by an interchange of l and m. Thus the same product appears with opposite sign, and they sum to zero.

It is interesting to note that if one considers more than two paths passing through the vertex k, then the fact that the alternating group contains exactly half of the members of the symmetric group can yield the same result.

The author would particularly welcome comments from readers. Further material can be supplied for those with such interest.

This paper results in part from joint work of the author and his friend and colleague Dragoš Cvetković of the University of Belgrade, whose textbook [3] presents linear algebra from a graph-theoretic viewpoint.

References

[1] C. L. Coates, Flow graph solutions of linear algebraic equations, IREE Trans. Circuit Theory, CT-6 (1959) 170–187.

[2] D. M. Cvetković, The determinant concept defined by means of graph theory, Mat. Vesnik, 12 (1975) 333–336.

[3] ______, Kombinatorina Teorija Matrica, Haunca Rnjiga, Belgrade, 1980.

[4] D. König, Uber Graphen und ihre Anwendung auf Determinantentheorie and Mengenlehre, Math. Ann., 77 (1916) 453–465.

[5] ______, Theorie der endlichen und unendlichen Graphen, Akadem. Verlagsges., Leipzig, 1936.

[6] S. Lang, Linear Algebra, Addison-Wesley, Reading, Mass., 1971.