A COMBINATORIAL APPROACH TO MATRIX ALGEBRA

Doron ZEILBERGER*

Department of Mathematics, Drexel University, Philadelphia, PA 19104, U.S.A.

Received 8 June 1984

"The theory of correspondence reaches far deeper than that of mere numerical congruity with which it is associated as the substance with the shadow"

James Joseph Sylvester

Introduction

To most contemporary mathematicians matrices and linear transformations are practically interchangeable notions. Indeed, the mainstream 'Bourbakian' establishment, with its profound disdain of the concrete, goes as far as to frown at the mere mention of the word 'matrix'.

To me, however, (as well as to a growing number of mathematical dissidents called 'combinatorialists') a matrix has nothing whatsoever to do with that intimidating abstract concept called 'a linear transformation between linear vector spaces'. Instead, an \(n \times n\) matrix is the 'blueprint' of all the possible edges one can draw on \(n\) given vertices, a determinant is the 'weight' of all permutation graphs and matrix-products represent paths (details later).

The purpose of this paper is to give a survey of this combinatorial interpretation of matrix algebra and to present elegant and illuminating proofs of five classical matrix identities.

In 1965, Dominque Foata [4, 2] gave a beautiful combinatorial proof of the celebrated MacMahon master theorem, thus setting the stage for combinatorial matrix algebra. Recently, two other elegant proofs have appeared: Straubing's proof of Cayley-Hamilton [9], and Orlin [8], Garsia [6] and Temperley [10] independently found a combinatorial proof of the matrix tree theorem.

I am going to present here new renditions of these three pearls, making them purely bijective and as succinct as possible. To them I am going to add two rubies of my own: a proof of \(\det(AB) = (\det A)(\det B)\) and a new combinatorial proof (quite shorter than Foata's [5]) of Jacobi's \(\det(e^A) = e^{tr^A}\).

* This research was partly supported by a summer research grant donated by my wife Jane.

1. The set-up

For us, the entries of matrices \( A = (a_{ij}) \) are not numbers but rather common indeterminates. We have \( n \) labeled vertices \( \{1, \ldots, n\} \) and the weight of the edge \( i \rightarrow j \) is \( a_{ij} \). A (directed) graph is a collection of edges and the weight of a graph is the product of the weights of its edges. For example,

\[
\text{weight} \left( \begin{array}{ccc}
1 & 2 & 4 \\
3 & 4 & \\
\end{array} \right) = a_{12}a_{13}a_{24}a_{34}.
\]

Whenever we have a set of objects possessing weights, we define the weight of set to be the sum of all the individual weights. For example,

\[
\text{weight} \left( \begin{array}{ccc}
2 & 3 & 1 \\
4 & 2 & 2 \\
3 & 1 & 3 \\
1 & 4 & 1 \\
\end{array} \right) = a_{12}a_{13} + a_{23}a_{31} + a_{34}.
\]

A cycle is a directed graph whose edges are \( i_1 \rightarrow i_2, i_2 \rightarrow i_3, \ldots, i_k \rightarrow i_1 \) some subset of the vertices \( \{i_1, \ldots, i_k\} \). The weight of a cycle is \( -a_{i_1i_2}a_{i_2i_3} \cdots \) (that is the negative of its weight qua graph). The weight of a disjoint union of cycles is defined as the product of the weights of all constituent cycles. In particular, it is readily seen that the weight of a permutation graph, whose edges are \( i \rightarrow \pi(i) \) \( (i = 1, \ldots, n) \) for some permutation \( \pi \) is equal to

\[
(-1)^{\# \text{cycles}} \prod_{i=1}^{n} a_{\pi(i)} = (\text{sgn } \pi) \prod_{i=1}^{n} (-a_{\pi(i)}).
\]

(This is so since the sign of an even cycle is \(-1\) and the sign of an even cycle is the sign of odd cycles, thus the sign of \( \pi \) is \((-1)^{\# \text{of even cycles}}\), taking \((-a_{ij})\) rather than \((a_{ij})\) gives a credit' to each odd cycle, making the total contribution to the left hand side of \((-1)^{\# \text{cycles}}\), as it should.)

We have thus obtained the following combinatorial interpretation of determinant;

\[
\text{det}(-a_{ij}) = \text{weight}(\mathcal{P}er(n)),
\]

where \( \mathcal{P}er(n) \) is the set of permutation graphs on the \( n \) vertices \( \{1, \ldots, n\} \). Similarly, the principal minors of \((-a_{ij})\) corresponding to any subset of vertices the weight of the set of disjoint unions of cycles covering these vertices. If \( \text{det}(\delta_{ij} - a_{ij}) \) (where \( \delta_{ij} \) is the identity matrix) is the weight of the set of all directed graphs that consist of disjoint union of cycles. For example, if \( n = 2 \)

\[
\text{det}(\delta_{ij} - a_{ij}) = \begin{vmatrix} 1-a_{11} & -a_{12} \\ -a_{21} & 1-a_{22} \end{vmatrix} = \text{weight} \left( \begin{array}{ccc} 1 & 2 \\ 1 & 2 \end{array} \right) + \text{weight} \left( \begin{array}{ccc} 1 & 1 \\ 1 & 2 \end{array} \right)
\]
A combinatorial approach to matrix algebra

\[ + \text{weight} \left( \begin{array}{c} 1 \\ \phi \end{array} \right) + \text{weight} \left( \begin{array}{c} \phi \\ 1 \end{array} \right) + \text{weight} \left( \begin{array}{c} \phi \\ \phi \end{array} \right) \]

If \( A = (a_{ij}) \) describes one kind of edges (called \( A \)-edges) and \( B = (b_{ij}) \) describes another kind of edges (called \( B \)-edges) then, for every pair \((i, j)\), the \((i, j)\) component of \( AB \) is the weight of the set of paths of length 2 from \( i \) to \( j \) such that the first edge is an \( A \)-edge and the second edge is a \( B \)-edge. This follows immediately from the definition of matrix multiplication. In particular the \((i, j)\) entry of \( A^k \) is the weight of the set of paths of length \( k \) from \( i \) to \( j \), where, of course,

\[ \text{weight}(i \rightarrow i_2 \rightarrow i_3 \rightarrow \cdots \rightarrow i_k \rightarrow j) = a_{ii}a_{i_2i_3}\cdots a_{i_kj}. \]

2. Foata's proof of the MacMahon master theorem \([2, 4]\)

Let \( A(m_1, \ldots, m_n) = \text{coefficient of } x_1^{m_1} \cdots x_n^{m_n} \text{ in } (a_{11}x_1 + \cdots + a_{1n}x_n)^{m_1} \cdots (a_{n1}x_1 + \cdots + a_{nn}x_n)^{m_n}. \) The MacMahon master theorem says that

\[ (\sum A(m_1, \ldots, m_n)x_1^{m_1} \cdots x_n^{m_n})\det(\delta_{ij} - a_{ij}) = 1. \quad (2) \]

Consider the collection \( \mathcal{A} \) of all pairs \((G, H)\) such that

(I) \( G \) is a directed graph, multiple edges and loops allowed such that

(i) For every vertex \( i \), the number of outgoing edges equals the number of incoming edges,

(ii) for every vertex \( i \), its outgoing edges are ordered from top to bottom (what computer folks would call a stack);

(II) \( H \) is a disjoint union of cycles (not necessarily covering all vertices).

For example, the following \((G, H)\) is such a pair:

\( G: \) out of 1: 1 \( \rightarrow \) 3 \quad \begin{align*} 
1 & \rightarrow 2 \\
1 & \rightarrow 1 \\
1 & \rightarrow 1 
\end{align*} \quad \begin{align*} 
\text{out of 2:} & \quad 2 \rightarrow 3 \\
 & \quad 2 \rightarrow 1 \\
 & \quad 2 \rightarrow 2 \end{align*} \quad \begin{align*} 
\text{out of 3:} & \quad 3 \rightarrow 3 \\
 & \quad 3 \rightarrow 3 \\
 & \quad 3 \rightarrow 1 \\
 & \quad 3 \rightarrow 2 
\end{align*} \quad \begin{align*} 
\text{H: (13) (i.e.,} & \quad 1 \rightarrow 3 \rightarrow 1) \\
\end{align*}
The weight of an edge \(i \rightarrow j\) is \(a_{ij}x_i\) and let
\[
\text{weight}(G, H) = (-1)^{\# \text{cycles of } H} \cdot \text{product of all edge-weights of } G.
\]

For example, for the above \((G, H)\)
\[
\text{weight}(G, H) = (-1)(a_{13}x_3)(a_{12}x_2)(a_{11}x_1) \\
\quad \cdot (a_{23}x_3)(a_{21}x_1)(a_{22}x_2) \\
\quad \cdot (a_{33}x_3)(a_{31}x_1)(a_{32}x_2) \\
\quad \cdot (a_{13}x_3)(a_{31}x_1) \\
= -a_{11}^2a_{12}a_{13}a_{21}a_{23}a_{31}a_{32}a_{33}x_1^3x_2^5x_3^5.
\]

We will prove (2) by showing that both the sides of (2) are equal to the same thing, namely to
\[
\text{weight}(\mathcal{A}) \overset{\text{def}}{=} \sum \text{weight}(G, H).
\]

Let \(\mathcal{G}\) be the set of all directed graphs satisfying (I) and \(\mathcal{H}\) the set of all such graphs satisfying (II). Clearly \(\mathcal{A} = G \times \mathcal{H}\) and
\[
\text{weight}(\mathcal{A}) = \text{weight}(\mathcal{G}) \cdot \text{weight}(\mathcal{H}).
\]

By the remarks in Section 1,
\[
\text{weight}(\mathcal{H}) = \det(\delta_{ij} - a_{ij}x_i).
\]

In order to show that the left-hand side of (2) is equal to weight \((\mathcal{A})\) we will prove that
\[
\text{weight}(\mathcal{G}) = \sum A(m_1, \ldots, m_n)x_1^{m_1} \cdots x_n^{m_n}.
\]

Indeed, for every \((m_1, \ldots, m_n)\) consider the subset of \(\mathcal{G}\) consisting of all graphs such that: for \(i = 1, \ldots, n\), \(i\) has \(m_i\) outgoing edges (and therefore \(m_i\) incoming edges). Now for every \(i\), you have \(m_i\) choices of choosing its outgoing edges, the total weight of each choice is \((a_{11}x_1 + \cdots + a_{m_i}x_{m_i})^m\) implying that the weight of all \(m_i\) choices outgoing edges of \(i\) is \((a_{11}x_1 + \cdots + a_{m_i}x_{m_i})^{m_i}\). Doing the same thing for every single vertex shows that the weight of the set of graphs having (for \(i = 1, \ldots, n\)) \(m_i\) edges out of \(i\) is \((a_{11}x_1 + \cdots + a_{m_i}x_{m_i})^{m_i} \cdots (a_{1n}x_1 + \cdots + a_{m_n}x_{m_n})^{m_n}\). But we also have to take care of the fact that there are exactly \(m_i\) edges coming into \(i\) (\(i = 1, \ldots, n\)) and therefore \(\text{weight}(\mathcal{G}) = x_1^{m_1} \cdots x_n^{m_n}\) in the above product = \(A(m_1, \ldots, m_n)x_1^{m_1} \cdots x_n^{m_n}\). Summing over \((m_1, \ldots, m_n)\) yields (5), which together with (4) and (3) yields
\[
\text{weight}(\mathcal{A}) = \text{left-hand side of (2)}.
\]

We will now prove that \(\text{weight}(\mathcal{A}) = 1\), and thus complete the proof. Let's define a mapping from \(\mathcal{A}\) to \(\mathcal{A}\) as follows.
Given a pair \((G, H)\), start at vertex 1 and walk along \(G\) in such a way that you always choose the top edge. Keep walking until either

*Case I.* You have encountered a previously visited vertex of \(G\), or
*Case II.* You have come across a vertex of \(H\).

In Case I we have transversed a complete cycle of \(G\) that is completely disjoint to the vertices of \(H\). We remove this cycle from \(G\) and put it in \(H\).

In Case II, we take the cycle of \(H\) to which that vertex belongs and move it from \(H\) to \(G\). Also, we do it in such a way that these newcomer edges of \(G\) are placed on the top of the old edges.

For example, if

\[
G = 1 \rightarrow 2 \quad H = \text{empty}
\]

then the walk on \(G\) is \(1 \rightarrow 2 \rightarrow 3 \rightarrow 2\), and Case I holds; thus the new \((G, H)\), call it \((G', H')\), is

\[
G' = 1 \rightarrow 2 \quad H' = (2, 3)
\]

then the walk on \(G\) is \(1 \rightarrow 2 \rightarrow 3 \rightarrow 2\), and Case I holds; thus the new \((G, H)\), call it \((G', H')\), is

\[
G' = 1 \rightarrow 2 \quad H' = (2, 3)
\]

Now let's apply the mapping to \((G', H')\). The walk is \(1 \rightarrow 2\), since vertex 2 is an \(H'\) vertex, belonging to \((2, 3)\). We remove \((2, 3)\) from \(H'\) and put its edges \(2 \rightarrow 3\) and \(3 \rightarrow 2\) in \(G'\) in their respective places on the top of the outgoing edges of 2 and 3 respectively. We get \((G, H)\) back. Of course this is no coincidence, and it is readily seen that applying the mapping twice on any pair \((G, H)\) reproduces it. In short, our mapping is an involution and therefore, of course, a bijection. Since there is 'conservation of edges' in \((G, H)\) the absolute value of the weight remains the same, but since the parity of the number of cycles of \(H\) changes, the sign changes. Thus all the terms of weight(\(\mathcal{A}\)) = \(\sum\) weight\((G, H)\) can be arranged in mutually cancelling pairs, except to the only element of \(\mathcal{A}\) on which the involution cannot be defined, namely the 'trivial' pair \((\text{empty}, \text{empty})\) whose weight is 1. Thus

weight(\(\mathcal{A}\)) = 1 = right-hand side of (2).

This completes the proof.

Let $A$ be an $n \times n$ matrix and let $P(\lambda) = \det(\lambda I - A)$ then the Cayley–Hamilton theorem says that the $n \times n$ matrix $P(A)$ is the zero matrix. Spelled out in a matrix says that

$$A^n + (-a_{11} - a_{22} - \cdots - a_{nn})A^{n-1}
+ \text{(sum of all } 2 \times 2 \text{ principal minors, of } -A)A^{n-2}
+ \cdots + \text{(sum of all } k \times k \text{ principal minors of } -A)A^{n-k}
+ \cdots + \det(-A) = 0.$$

We have to prove that every entry of the matrix on the left hand side of (*) equal to zero.

Fix $i$ and $j$ and let $\mathcal{A} = \mathcal{A}(i, j)$ be the set of pairs $(P, C)$ such that

(i) $P$ is a path from $i$ to $j$,
(ii) $C$ is a disjoint union of cycles,
(iii) The total number of edges of $P$ and $C$ combined equals $n$.

The weight of an edge $k \to m$ is $a_{km}$ and

$$\text{weight}(P, C) = (-1)^{\# \text{cycles of } C} \prod \text{edge-weights of } C.$$

For example, if $i = 1, j = 2, n = 5, (1\rightarrow 3 \rightarrow 2, (1)(3,5))$ is an element of whose weight is $(-1)^2(a_{13}a_{32})(a_{11})(a_{35}a_{53})$.

Now we claim that

$$\text{weight}(\mathcal{A}(i, j)) = (i, j) \text{ entry of the left-hand side of (*)}.$$

Indeed, the path $P$ may be of any length $n-k$ for $0 \leq k \leq n$. The weight of the $k$ paths of length $n-k$ from $i$ to $j$ is exactly the $(i, j)$ entry of $A^{n-k}$. Now we have $k$ edges left to form disjoint cycles, and you have the freedom to choose a $k$-element subset of $\{1, \ldots, n\}$ for your vertices. The weight of the set of all this is (by the remarks of Section 1) equal to the sum of all $k \times k$ principal minors of $-A$. Summing over all $0 \leq k \leq n$ gives (7).

The proof will be completed once we show that for every $i, j$

$$\text{weight}(\mathcal{A}(i, j)) = 0.$$

To this end we will introduce the following mapping from $\mathcal{A}(i, j)$ to itself. Given $$(P, C)$ start at $i$ and walk along the path $P$ until you either

Case I. Come to a previously visited vertex of $P$, or
Case II. come to a vertex that belongs to one of the cycles of $C$.

In Case I you have transversed a cycle of $P$ whose vertices are disjoint to all the cycles of $C$. You remove that cycle from $P$ and join it to $C$.

In Case II you remove that cycle from $C$ and insert it (at that vertex) in $P$. 
Example. \( n = 5, \ i = 1, \ j = 3 \)

\[
(1 \to 2 \to 3 \to 2 \to 3; (5)) \leftrightarrow (1 \to 2 \to 3; (23), (5))
\]

\[
(1 \to 3 \to 3; (3, 4, 5)) \leftrightarrow (1 \to 3 \to 4 \to 5 \to 3 \to 3; \emptyset).
\]

It is readily seen that this mapping is an involution defined on every element \((P, C)\) of \(\mathcal{A}\). (Let the number of vertices (= number of edges) of \(C\) be \(k\), and suppose that the vertices of \(P\) are disjoint from those of \(C\). Then \(P\) has as many vertices as edges \((n-k\) of them) and therefore must contain a cycle.) By 'conservation of edges' the absolute value of the weight stays the same, but since the parity of the number of cycles of \(C\) changes, the sign of the weight is reversed. Thus, all the elements of \(\text{weight}(\mathcal{A})\) can be arranged in mutually cancelling pairs and their sum is therefore zero.

4. A combinatorial proof of the matrix tree theorem [6, 8, 10]

Consider directed graphs on the \(n\) vertices \(\{1, \ldots, n\}\). A tree rooted at \(n\) is a directed graph without cycles such that every vertex has exactly one outgoing edge except to the root \(n\) that has no outgoing edges. Let \(\mathcal{T} = \mathcal{T}(n)\) be the set of trees rooted at \(n\). The weight of an edge \(k \to m\) is \(a_{km}\) and the weight of a tree (or for that matter any directed graph) is the product of the edge-weights.

The matrix-tree theorem says that \(\text{weight}(\mathcal{T}(n))\) equals the determinant

\[
\begin{vmatrix}
   a_{12} + \cdots + a_{1n} & -a_{12} & \cdots & -a_{1,n-1} \\
   -a_{21} & a_{21} + \cdots + a_{2n} & & -a_{2,n-1} \\
   & & \ddots & \\
   -a_{n-1,1} & -a_{n-1,2} & \cdots & a_{n-1,1} + \cdots + a_{n-1,n}
\end{vmatrix}
\]  \tag{9}

Let \(\mathcal{B}\) be the set of pairs \((B, C)\) such that

(i) \(B\) is a directed graph such that for a certain subset \(V_B\) of \([1, \ldots, n-1]\) there is exactly one edge going out of every vertex of \(V_B\). The end vertex of each edge may be any vertex of \([1, \ldots, n]\) except its origin (i.e., no slings allowed);

(ii) \(C\) is a collection of disjoint cycles, of length \(\geq 2\), on the set of vertices \(V_C\), \(V_C\) being the complement of \(V_B\) with respect to \([1, \ldots, n-1]\).

The weight of a pair \((B, C)\) is defined by

\[
\text{weight}(B, C) = (-1)^{\#\text{cycles of } C}[\text{product of all edge-weights of } B \text{ and } C].
\]

For example \((n = 5)\)

\[
\text{weight}(1 \to 5, 3 \to 5; (2, 4)) = (-1)^1 a_{15} a_{35} a_{24} a_{42}.
\]

It is readily seen that \((9) = \text{weight}(\mathcal{B})\).

Define the following mapping on \(\mathcal{B}\). Given \((B, C)\) look at all cycles, both of \(B\) (if
any) and of \( C \). Pick the cycle that contains the lowest vertex and change affiliation (if it belonged to \( B \) put it in \( C \) and vice versa). For example \((n = 6)\):

\[
(1 \rightarrow 2, 2 \rightarrow 1, 4 \rightarrow 6; (35)) \leftrightarrow (4 \rightarrow 6; (12)(35))
\]

\[
(1 \rightarrow 6, 2 \rightarrow 6; (345)) \leftrightarrow (1 \rightarrow 6, 2 \rightarrow 6, 3 \rightarrow 4, 4 \rightarrow 5, 5 \rightarrow 3; \emptyset).
\]

It is not hard to see that we have a sign reversing involution that is defined on all elements of \( \mathcal{B} \) that have cycles. The only survivors are those elements of \( \mathcal{B} \) of the form \((B, \emptyset)\) where \( B \) has no cycles, i.e., is a tree! Thus \( \text{weight}(\mathcal{A}) = \text{weight}(\mathcal{T}) \) and this completes the proof that \( \text{weight}(\mathcal{F}) \) equals \((9)\).

5. \( \det(AB) = (\det A)(\det B) \)

The matrix \( AB \) represents compound edges \( i \xrightarrow{A} k \xrightarrow{B} j \) with weight \( a_{jk}b_{kj} \), where \( k \) can be any vertex. Let \( \text{weight}_A(\pi) = (\text{sgn } \pi)a_{1\pi(1)} \cdots a_{n\pi(n)} \), \( \text{weight}_B(\pi) = (\text{sgn } \pi)b_{1\pi(1)} \cdots b_{n\pi(n)} \). Let \( \text{Per}(n) \) denote the set of permutations on \( \{1, \ldots, n\} \), then \( \det A = \text{weight}_A(\text{Per}(n)) \), \( \det B = \text{weight}_B(\text{Per}(n)) \). What is \( \det(AB) \)?

Let \( Z(n) \) be the set of pairs \((f, \pi)\) where \( f \) is any mapping \( \{1, \ldots, n\} \rightarrow \{1, \ldots, n\} \) and \( \pi \) is a permutation. Let \( \text{weight}(f, \pi) = (\text{sgn } \pi)(a_{1f(1)}b_{f(1)f(1)}a_{2f(2)}b_{f(2)f(2)} \cdots a_{nf(n)}b_{f(n)f(n)}) \).

A moment’s reflection would convince you that

\[
\det(AB) = \text{weight}(Z(n)).
\]

An element of \( Z(n) \) is a good guy if \( f \) is a permutation. Then of course \( f^{-1} \) is a permutation and \( \text{weight}(f, \pi) = \text{weight}_A(f)\text{weight}_B(f^{-1} \circ \pi) \). Thus

\[
\sum_{(f, \pi) \text{ good}} \text{weight}(f, \pi) = (\det A)(\det B). \tag{11}
\]

In order to prove that \( \det(AB) \), which we said was equal to \( \text{weight}(Z(n)) \), is equal to \( (\det A)(\det B) \), we have to show only, thanks to (10), that

\[
\sum_{(f, \pi) \text{ bad}} \text{weight}(f, \pi) = 0. \tag{11}
\]

Once again we have to find a killer involution. If \((f, \pi)\) is a bad guy, all it means is that \( f \) is not a permutation, i.e., there exist \( b, i \) and \( i' \) such that \( f(i) = b \) and \( f(i') = i' \). Let \( f(i) = b \) and \( f(i') = i' \) in a more picturesque notation there exist \( A \)-edges \( i \xrightarrow{A} b \) and \( i' \xrightarrow{A} b \). Pick the smallest such \( b \), and for that \( b \), the smallest such \( i \) and \( i' \).

Case I: \( i \) and \( i' \) belong to the same cycle of \( \pi \). The cycle to which both \( i \) and \( i' \) belong looks as follows:

\[
i \xrightarrow{A} b \xrightarrow{B} \pi(i) \xrightarrow{A} \text{whatever} \cdots \xrightarrow{B} i' \xrightarrow{A} b \xrightarrow{B} \pi(i') \rightarrow \text{blablabla} \cdots \rightarrow i.
\]

What you have to do is break this long cycle into two cycles:

\[
i \xrightarrow{A} b \xrightarrow{B} \pi(i') \rightarrow \text{blablabla} \rightarrow i
\]
and
\[ \pi(i) \xrightarrow{B} \text{whatever } B \xrightarrow{b} B \xrightarrow{\pi(i)}. \]

(Note that the underlying permutation changed from \( \pi \) to \( \pi \) times the transposition \((i, i')\).)

Case II: \( i \) and \( i' \) belong to different cycles of \( \pi \). Let these cycles be
\[ i \xrightarrow{A} b \xrightarrow{B} \pi(i) \xrightarrow{A} \text{blablabla} \ldots B \xrightarrow{b} i \]
and
\[ i' \xrightarrow{A} b \xrightarrow{B} \pi(i') \xrightarrow{A} \text{whatever} \ldots B \xrightarrow{b} i'. \]

In this case what you have to do is to combine them into one cycle:
\[ i \xrightarrow{A} b \xrightarrow{B} \pi(i') \xrightarrow{A} \text{whatever} \ldots B \xrightarrow{b} \pi(i) \rightarrow \text{blablabla} \ldots \rightarrow i. \]

(Note that the underlying permutation changed from \( \pi \) to \( \pi \) times the transposition \((i, i')\).)

Example. \( n = 6 \)

\[ (1 \xrightarrow{A} 4 \xrightarrow{B} 2 \xrightarrow{A} 2 \xrightarrow{B} 5 \xrightarrow{A} 3 \xrightarrow{B} 6 \xrightarrow{A} 3 \xrightarrow{B} 2 \xrightarrow{A} 4 \xrightarrow{B} 2 \xrightarrow{A} 1) \]
\[ \cong \]
\[ (1 \xrightarrow{A} 4 \xrightarrow{B} 2 \xrightarrow{A} 2 \xrightarrow{B} 4 \xrightarrow{A} 2 \xrightarrow{B} 1)(2 \xrightarrow{B} 5 \xrightarrow{A} 3 \xrightarrow{B} 6 \xrightarrow{A} 3 \xrightarrow{B} 3 \xrightarrow{A} 2). \]

It is readily seen that what we have here is a sign reversing involution defined on all the bad guys and thus the sum of the weights of all the bad guys is 0. This proves (11) which together with (10) completes the proof of \( \det(AB) = (\det A)(\det B) \).

6. A new combinatorial proof of Jacobi's \( \det(e^A) = e^{\pi A} \)

The first to realize that Jacobi's identity has anything to do with combinatorics was Jackson [7] who gave it a combinatorial interpretation. Foata [5] then went on to give an elegant combinatorial proof. We are now going to give another combinatorial proof that is shorter and more direct.

\[ e^A = \sum A^k/k! \] is the exponential generating function of paths of all length. Namely, writing \( B = e^A, B = (b_{ij}) \) we have

\[ \frac{1}{k!} \text{weight[set of all paths from } i \text{ to } j \text{ of length } k \} \]

= the sum of all terms in \( b_{ij} \) of total degree \( k \).

Now for \( m = 0, 1, 2, \ldots \), let \( \mathcal{B}_m \) be the set of objects of the form \( (\pi, P_{\pi(i)}), \)
\(i = 1, \ldots, n\) where

(1) \(\pi\) is a permutation of \(\{1, \ldots, n\}\);
(2) For \(i = 1, \ldots, n\), \(P_{\pi(i)}\) is a path from \(i\) to \(\pi(i)\);
(3) The total number of edges of all paths is \(m\);
(4) The edges are labeled by distinct labels from \(\{1, \ldots, m\}\) in such a way that they are increasing along every path.

The weight of such an object is \(\text{sgn} \ \pi\) times the product of all edge-weights, the product of an edge \(k \rightarrow l\) being \(a_{kl}\).

For example if \(n = 4\), \(m = 15\)

\[
\pi = 1 \rightarrow 2, 2 \rightarrow 1, 3 \rightarrow 3, 4 \rightarrow 4
\]

\(P_{12}: 1 \rightarrow 2 \rightarrow 3 \rightarrow 5 \rightarrow 2\)

\(P_{21}: 2 \rightarrow 4 \rightarrow 1 \rightarrow 9 \rightarrow 4 \rightarrow 1\)

\(P_{33}: 3 \rightarrow 4 \rightarrow 8 \rightarrow 1 \rightarrow 10 \rightarrow 2 \rightarrow 14 \rightarrow 3\)

\(P_{44}: 4 \rightarrow 7 \rightarrow 4 \rightarrow 1 \rightarrow 13 \rightarrow 4 \rightarrow 15 \rightarrow 4\)

is one member of \(\mathcal{B}_{15}\) whose weight is

\[(+1)(a_{12}a_{22}a_{22})(a_{21}a_{11}a_{14}a_{41})(a_{34}a_{41}a_{12}a_{23})(a_{44}a_{41}a_{14}a_{44}).\]

By general properties of exponential generating functions we have

\[\text{det}(e^A) = \sum_{m=0}^{\infty} \frac{1}{m!} \text{weight} (\mathcal{B}_m).\]

We are now going to define an involution \(\mathcal{B}_m \rightarrow \mathcal{B}_m\) (for every \(m\)) that is going to get rid of most of the terms in weight(\(\mathcal{B}_m\)).

Let \(j \rightarrow i\) be the edge of highest label \(s\) for which \(j \neq i\). This edge must necessarily belong to \(P_{\pi^{-1}(i)}\) which has the form

- first path:

\(P_{\pi^{-1}(i)}: \pi^{-1}(i) \rightarrow \text{whatever} \rightarrow j \rightarrow i \rightarrow \cdot \cdot \cdot \rightarrow i\)

where \(s = s_0\) and \(r \geq 0\).

Now consider \(P_{\pi^{-1}(j)}\)

- second path:

\(P_{\pi^{-1}(j)}: \pi^{-1}(j) \rightarrow \text{blabla} \rightarrow \cdot \cdot \cdot \rightarrow j \rightarrow i \rightarrow \cdot \cdot \cdot \rightarrow i\)

Let \(0 \leq \alpha \leq l\) be the only \(\alpha\) such that \(t_\alpha < s < t_{\alpha+1}\). The involution consists of swapping the portion \(i \rightarrow i \rightarrow \cdot \cdot \cdot \rightarrow i\) of the first path and (the possibility?
A combinatorial approach to matrix algebra

empty) portion \( \xrightarrow{i_n-1} j \xrightarrow{i_n-2} j \cdots \xrightarrow{i_1} j \) of the second path, getting for the transformed object

- first path:
  \[ P_{\pi^{-1}(i)_1} : \pi^{-1}(i) \longrightarrow \text{whatever} \cdots \longrightarrow j \xrightarrow{j^*} j \longrightarrow \cdots \xrightarrow{j_1} j \]

- second path:
  \[ P_{\pi^{-1}(i)_2} : \pi^{-1}(j) \longrightarrow \text{blablabla} \cdots \xrightarrow{j^*} j \xrightarrow{j^*} i \xrightarrow{j^*} i \longrightarrow \cdots \xrightarrow{j_1} i \]

**Example.** \( n = 3 \)

\[
\begin{align*}
\pi &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \\
\pi_{12} &= \begin{pmatrix} 1 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 6 \rightarrow 2 \rightarrow 13 \rightarrow 2 \\
P_{12} &= \begin{pmatrix} 1 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 6 \rightarrow 2 \rightarrow 13 \rightarrow 2 \\
P_{12} &= \begin{pmatrix} 1 \rightarrow 2 \rightarrow 3 \rightarrow 2 \rightarrow 6 \rightarrow 2 \rightarrow 13 \rightarrow 2 \\
P_{23} &= \begin{pmatrix} 3 \rightarrow 4 \rightarrow 1 \rightarrow 5 \rightarrow 1 \rightarrow 11 \rightarrow 1 \rightarrow 12 \rightarrow 1 \\
P_{23} &= \begin{pmatrix} 3 \rightarrow 4 \rightarrow 1 \rightarrow 5 \rightarrow 1 \rightarrow 11 \rightarrow 1 \rightarrow 12 \rightarrow 1 \\
P_{31} &= \begin{pmatrix} 3 \rightarrow 4 \rightarrow 1 \rightarrow 5 \rightarrow 1 \rightarrow 11 \rightarrow 1 \rightarrow 12 \rightarrow 1 \\
P_{31} &= \begin{pmatrix} 3 \rightarrow 4 \rightarrow 1 \rightarrow 5 \rightarrow 1 \rightarrow 11 \rightarrow 1 \rightarrow 12 \rightarrow 1 \\
\end{align*}
\]

Since we have 'conservation of edges', the absolute value of the weight is retained, but the sign changes, since the underlying permutation has been multiplied by the transposition \((i, j)\). Thus the involution gets rid of all the terms in weight\((\mathcal{B}_m)\) corresponding to elements of \(\mathcal{B}_m\) on which the involution makes sense. The only survivors of weight\((\mathcal{B}_m)\) are weights of elements on which the involution cannot be defined, call these elements \(\mathcal{C}_m\). Thus weight\((\mathcal{B}_m) = \text{weight}(\mathcal{C}_m)\), where \(\mathcal{C}_m\) are those objects all of whose edges are slings, that is, edges from a vertex to itself \(i \rightarrow i\). The underlying permutation for all the members of \(\mathcal{C}_m\) must necessarily be the identity permutation and all paths have the form

\[ P_{11} : 1 \rightarrow 1 \rightarrow 1 \rightarrow \cdots \rightarrow 1 \]

\[ P_{nn} : n \rightarrow n \rightarrow \cdots \rightarrow n \]

with appropriate edge labels. Since the exponential generating function for paths of the form \(i \rightarrow i \rightarrow \cdots \rightarrow i\) is \(e^a\) it follows that

\[
\sum_{m=1}^{\infty} \frac{1}{m!} \text{weight}(\mathcal{C}_m) = e^a e^{a^2} \cdots e^{a^n} = e^{trA}.
\]

Thus

\[
\det(e^A) = \sum \frac{1}{m!} \text{weight}(\mathcal{B}_m) = \sum \frac{1}{m!} \text{weight}(\mathcal{C}_m) = e^{trA}
\]

This completes the proof.
7. Exercises

The difficulty rating of these exercises follows the famous Knuth ranges from 0 (outright trivial) to 50 (impossible).

1. (10) Starting from the combinatorial definition of the determinant, the determinant vanishes if two rows are identical.

2. (12) Give a combinatorial interpretation to an arbitrary minor (not necessarily principal) obtained by choosing rows $i_1, \ldots, i_k$ and columns $j_1, \ldots, j_k$.

3. (20) Using Exercise 2, prove the Laplace expansion of the determinant.

4. (39) Go through Aitken's book [1] and try to prove combinatorial results as possible.

5. (28) Using a proof similar to Straubing's, prove the following identity: for any integer $m$ if $A$ is an $n \times n$ matrix then if $A$ is an $n \times n$ matrix

$$(\text{tr } A^m) + (\text{sum of } 1 \times 1 \text{ principal minors of } -A) (\text{tr } A^{-1}) + \cdots + (\text{sum of } k \times k \text{ principal minors of } -A) (\text{tr } A^{m-k}) + \cdots + (\text{sum of } m \times m \text{ principal minors of } -A) m = 0.$$ 

Note that for $m \geq n$ it is a trivial consequence of the Cayley-Hamilton theorem and that for $A$ diagonal these are Newton's identities.

6. (45) Another way of stating the matrix tree theorem is to say that the number of the set of trees rooted at $n$ equals the $(n, n)$ minor of the determinant matrix $A_{ij}$ defined by $A_{ii} = \sum_{i \neq i} a_{ii}, A_{ij} = -a_{ij}$ ($i \neq j$). Find what is enumerated by an arbitrary (not necessarily principal) minor and prove the so-called "all minors matrix tree theorem". (For a proof see Chaiken's [3] interesting paper.)

References