A determinantal identity for the permanent of a rank 2 matrix

Adam W. Marcus*
Princeton University

August 30, 2016

Abstract

We prove a relationship between the permanent of a rank 2 matrix and the determinants of its Hadamard powers. The proof uses a decomposition of the determinant in terms of Schur polynomials.

Let $\mathbb{F}$ be a field. For a matrix $M \in \mathbb{F}^{n\times n}$ with entries $m(i,j)$ and integer p, let M_p be the matrix with

$$M_p(i,j) = m(i,j)^p.$$

When M has rank at most 2 and no nonzero entries, Carlitz and Levine showed [1]:

$$\det[M_{-2}] = \det[M_{-1}] \ \operatorname{perm}[M_{-1}] \tag{1}$$

where

$$\operatorname{perm}[M] = \sum_{\sigma\in S_n}\prod_{i=1}^{n} m(i,\sigma(i)) \qquad \text{and} \qquad \det[M] = \sum_{\sigma\in S_n}(-1)^{|\sigma|}\prod_{i=1}^{n} m(i,\sigma(i))$$

are the usual *permanent* and *determinant* matrix functions. The proof in [1] is elementary, using little more than the definitions and some facts concerning the cycle structure of permutations. In this paper we prove an identity in a similar spirit, but using an entirely different means. Our main result (Theorem 2.5) shows that when M has rank at most 2,

$$(n!)^2 \det[M_n] = (n^n)\det[M_{n-1}] \ \operatorname{perm}[M_1]. \tag{2}$$

The proof uses a decomposition of the determinant into Schur polynomials (Lemma 2.1).

1 Preliminaries

For a set S and a function f, we will write

$$a^S := \prod_{i\in S} a_i \qquad \text{and} \qquad f(S) = \{f(i)\}_{i\in S}.$$

We use the customary notation that $[n] = \{1,2,\dots,n\}$ and that $\binom{[n]}{k}$ denotes the collection of subsets of $[n]$ size k. For a permutation σ, we write $|\sigma|$ to denote the number of cycles in its cycle decomposition.

*Research supported by NSF CAREER grant 1552520.

1.1 Alternating Polynomials

We will say that a polynomial $p(x_1, \dots, x_n) \in \mathbb{F}[x_1, \dots, x_n]$ is *symmetric* if

$$p(x_1, \dots x_i, x_{i+1}, \dots, x_n) = p(x_1, \dots x_{i+1}, x_i, \dots, x_n)$$

and *alternating* if

$$p(x_1, \dots x_i, x_{i+1}, \dots, x_n) = -p(x_1, \dots x_{i+1}, x_i, \dots, x_n)$$

for all transpositions $(i, i+1)$. Since the set of transpositions generates the symmetric group, an equivalent definition is (for symmetric polynomials)

$$p(x_1, \dots, x_n) = p(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

and (for alternating polynomials)

$$p(x_1, \dots, x_n) = (-1)^{|\sigma|} p(x_{\sigma(1)}, \dots, x_{\sigma(n)}) \tag{3}$$

for all $\sigma \in S_n$. In particular, (3) implies that any alternating polynomial p must be 0 whenever $x_i = x_j$ for some $i \neq j$. One example of an alternating polynomial is the Vandermonde polynomial

$$\Delta(x_1, \dots, x_n) = \prod_{i<j} (x_i - x_j). \tag{4}$$

It is easy to see that the Vandermonde polynomial is an essential part of all alternating polynomials:

Lemma 1.1. *For all alternating polynomials $f(x_1, \dots, x_n)$, there exists a symmetric polynomial $t(x_1, \dots, x_n)$ such that*

$$f(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) t(x_1, \dots, x_n).$$

Proof. For distinct $y_2, \dots, y_n \in \mathbb{F}$, (3) implies that the univariate polynomial

$$g(x) = f(x, y_2, \dots, y_n) \in \mathbb{F}[x]$$

satisfies $g(y_k) = 0$ for each $k = 2, \dots, n$. Hence $(x - y_k)$ must be a factor of g and so $(x_1 - x_k)$ must be a factor of g. Since this is true for all k and all i (not just $i = 1$), every polynomial of the form $(x_i - x_k)$ must be a factor of f, and so

$$f(x_1, \dots, x_n) = \Delta(x_1, \dots, x_n) t(x_1, \dots, x_n)$$

for some polynomial t. To see that t is symmetric, one merely needs to note that

$$f(x_1, \dots, x_n) = (-1)^{|\sigma|} f(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

now implies

$$\Delta(x_1, \dots, x_n) t(x_1, \dots, x_n) = (-1)^{|\sigma|} \Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)}) t(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

where

$$\Delta(x_1, \dots, x_n) = (-1)^{|\sigma|} \Delta(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

and so

$$t(x_1, \dots, x_n) = t(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

$\square$

1.2 Young Tableaux

For a positive integer n, we will say that a sequence of nonnegative integers $\lambda = (\lambda_1, \lambda_2, \dots)$ is a *partition of* n if

1. $\lambda_i \geq \lambda_{i+1}$ for all $i \geq 1$
2. $\sum_i \lambda_i = n$

and we will write $|\lambda| = n$. We will let Λ_n denote the collection of partitions of n and $\Lambda_* = \bigcup_n \Lambda_n$. The *length* of a partition, written $\ell(\lambda)$, is the largest k for which $\lambda_k \neq 0$ (it should be clear from the definition that only a finite number of elements of λ_i can be nonzero, and that they must occupy an initial interval of λ). For ease of reading, we will use the customary exponential notation: that is, we will write the partition

$$(\underbrace{t_1, \dots, t_1}_{n_1 \text{ times}}, \underbrace{t_2, \dots, t_2}_{n_2 \text{ times}}, \dots)$$

as $(t_1^{n_1}, t_2^{n_2}, \dots)$. The one exception will be the values of 0, which we will never include in any presentation unless necessary (but which will always exist).

A *Young diagram* is a finite collection of boxes (called cells) which are arranged in left-justified rows with nonincreasing lengths. There is a natural bijection between partitions and Young diagrams where the ith row of the Young diagram has λ_i cells in it. A *Young tableau* is obtained by filling in the boxes of the Young diagram with positive integers (possibly restricted to some ground set Ω). A tableau is called *standard* (resp. *semistandard*) if the entries in each row and each column are strictly (resp. weakly) increasing. The sequence of integers $w(T) = \{t_i\}_{i=1}^{\infty}$ where t_i is the number of times the integer i appears in a given tableau is called the *weight sequence*.

1.3 Schur Polynomials

The degree d Schur polynomials in n variables form a linear basis for the space of homogeneous degree d symmetric polynomials in n variables, indexed by partitions λ with $|\lambda| = d$. For a partition λ, the Schur polynomial s_λ is defined as

$$s_\lambda(x_1, x_2, \dots, x_n) = \sum_T x^{w(T)} = \sum_T x_1^{t_1} \cdots x_n^{t_n}$$

where the summation is over all semistandard Young tableaux T of shape λ using ground set $\Omega = \{1, \dots, n\}$ and where $w(T) = t_1, \dots, t_n$ is the weight sequence of T.

Jacobi gave a more direct formula for computing Schur polynomials [2]: given a partition λ, define the functions

$$a_\lambda(x_1, x_2, \dots, x_n) = \det \begin{bmatrix} x_1^{\lambda_1+n-1} & x_2^{\lambda_1+n-1} & \dots & x_n^{\lambda_1+n-1} \\ x_1^{\lambda_2+n-2} & x_2^{\lambda_2+n-2} & \dots & x_n^{\lambda_2+n-2} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{\lambda_n} & x_2^{\lambda_n} & \dots & x_n^{\lambda_n} \end{bmatrix}.$$

In particular, when $\lambda = (0)$, the matrix involved is the well-known *Vandermonde matrix*, and so

$$a_{(0)}(x_1, x_2, \ldots, x_n) = \det \begin{bmatrix} x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \\ x_1^{n-2} & x_2^{n-2} & \ldots & x_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \ldots & 1 \end{bmatrix} = \Delta(x_1, \ldots, x_n).$$

where Δ is the alternating polynomial from (4). Since determinants are alternating with respect to transposition of columns, the polynomials a_λ are alternating with respect to transposition of its variables. Hence by Lemma 1.1, we can write

$$a_\lambda(x_1, \ldots, x_n) = a_{(0)}(x_1, \ldots, x_n) s_\lambda(x_1, \ldots, x_n) \tag{5}$$

where each $s_\lambda(x_1, \ldots, x_n)$ is a symmetric polynomial.

Theorem 1.2 (Jacobi)**.** *For all λ, the polynomials s_λ defined in (5) are the Schur polynomials.*

Of particular importance for us is that the Schur polynomials indexed by the partitions $\lambda = (1^k)$ are the elementary symmetric polynomials.

Lemma 1.3. *For all integers $n \geq 1$ and $k \geq 0$*

$$s_{(1^k)}(x_1, \ldots, x_n) = e_k(x_1, \ldots, x_n)$$

for all k. In particular, $s_{(0)}(x_1, \ldots, x_n) = 1$ (the case $k = 0$).

We end by noting that we will only use some of the most basic properties of Schur polynomials. They appear naturally in a variety of settings; for example, combinatorics (as a basis for symmetric functions), random matrix theory (as zonal spherical polynomials), and representation theory (as characters related to representations of the general linear groups). For a more comprehensive treatment of Schur polynomials, see [3].

2 Theorem

For this section, we fix vectors $a, b \in \mathbb{F}^n$ and define the matrices A_p with entries

$$A_p(i, j) = (1 + a_i b_j)^p.$$

Lemma 2.1. *For integer $p > 0$, we have*

$$\det[A_p] = \Delta(a)\Delta(b) \sum_{\substack{S \subseteq \{0, \ldots, p\} \\ |S| = n}} s_\lambda(a) s_\lambda(b) \left(\prod_{i \in S} \binom{p}{i} \right)$$

where $\lambda_j = S_{n-j+1} - n + j$.

Proof. By definition,

$$\begin{aligned}
\det[A_p] &= \sum_{\sigma\in S_n}(-1)^{|\sigma|}\prod_{j=1}^{n}A_p(j,\sigma(j))\\
&= \sum_{\sigma\in S_n}(-1)^{|\sigma|}\prod_{j=1}^{n}(1+a_ib_{\sigma(j)})^p\\
&= \sum_{\sigma\in S_n}(-1)^{|\sigma|}\prod_{j=1}^{n}\left(\sum_{i_j=0}^{p}\binom{p}{i_j}(a_jb_{\sigma(j)})^{i_j}\right)\\
&= \sum_{i_1,\dots,i_n=0}^{p}\binom{p}{i_1}\cdots\binom{p}{i_n}a_1^{i_1}\dots a_n^{i_n}\sum_{\sigma\in S_n}(-1)^{|\sigma|}b_{\sigma(1)}^{i_1}\cdots b_{\sigma(n)}^{i_n}
\end{aligned}$$

For a vector $\vec{i}=i_1,\dots,i_n$, let

$$f(\vec{i})=\sum_{\sigma\in S_n}(-1)^{|\sigma|}b_{\sigma(1)}^{i_1}\cdots b_{\sigma(n)}^{i_n}.$$

We first claim that $f(\vec{i})=0$ whenever $i_j=i_k$ for some $j\neq k$. To see that this is true, consider the matrix $W_{\vec{i}}(s,t)=b_t^{i_s}$. Then

$$\det\left[W_{\vec{i}}\right]=\sum_{\sigma\in S_n}(-1)^{|\sigma|}\prod_{s=1}^{n}W_{\vec{i}}(s,\sigma(s))=\sum_{\sigma\in S_n}(-1)^{|\sigma|}\prod_{j=1}^{n}b_{\sigma(s)}^{i_s}=f(\vec{i}).$$

But if $i_j=i_k$ for $j\neq k$, then W has rows j and k the same (and so this determinant is 0). Hence we have

$$\begin{aligned}
\det[A_p] &= \sum_{i_1\neq\cdots\neq i_n=0}^{p}\binom{p}{i_1}\cdots\binom{p}{i_n}a_1^{i_1}\dots a_n^{i_n}\sum_{\sigma\in S_n}(-1)^{|\sigma|}b_{\sigma(1)}^{i_1}\cdots b_{\sigma(n)}^{i_n}\\
&= \sum_{\pi\in S_n}\sum_{0\le i_1<\cdots<i_n\le p}\binom{p}{i_1}\cdots\binom{p}{i_n}a_1^{i_{\pi(1)}}\dots a_n^{i_{\pi(n)}}\sum_{\sigma\in S_n}(-1)^{|\sigma|}b_{\sigma(1)}^{i_{\pi(1)}}\cdots b_{\sigma(n)}^{i_{\pi(n)}}\\
&= \sum_{\pi\in S_n}\sum_{0\le i_1<\cdots<i_n\le p}\binom{p}{\pi(i_1)}\cdots\binom{p}{\pi(i_n)}a_{\pi(1)}^{i_1}\dots a_{\pi(n)}^{i_n}\sum_{\sigma\in S_n}(-1)^{|\sigma|}b_{\sigma(\pi(1))}^{i_1}\cdots b_{\sigma(\pi(n))}^{i_n}\\
&= \sum_{0\le i_1<\cdots<i_n\le p}\binom{p}{i_1}\cdots\binom{p}{i_n}\sum_{\pi,\sigma\in S_n}(-1)^{|\sigma|+|\pi|}a_{\pi(1)}^{i_n}\dots a_{\pi(n)}^{i_1}b_{\sigma(1)}^{i_n}\cdots b_{\sigma(n)}^{i_1}\\
&= \Delta(a)\Delta(b)\sum_{\substack{S\subseteq\{0,\dots,p\}\\|S|=n}}s_\lambda(a)s_\lambda(b)\left(\prod_{i\in S}\binom{p}{i}\right)
\end{aligned}$$

where one can calculate that the appropriate $\lambda_j=S_{n-j+1}-n+j$. $\square$

Corollary 2.2.

$$\det[A_{n-1}]=\Delta(a)\Delta(b)\left(\prod_{j=0}^{n-1}\binom{n-1}{j}\right)$$

Proof. By Lemma 2.1, we have

$$\det\left[A_{n-1}\right] = \Delta(a)\Delta(b) \sum_{\substack{S\subseteq\{0,\ldots,n-1\}\\|S|=n}} s_\lambda(a)s_\lambda(b)\left(\prod_{i\in S}\binom{p}{i}\right)$$

but the only set satisfying the constraint in the summation is the set $\{0,\ldots,n-1\}$ itself. One can check that this leads to having $\lambda=(0)$, which by Lemma 1.3 make $s_\lambda(a)s_\lambda(b)=1$. □

Lemma 2.3.

$$\operatorname{perm}\left[A\right] = \sum_{k=0}^{n} k!(n-k)!e_k(a)e_k(b)$$

Proof. By definition, we have

$$\operatorname{perm}\left[A\right] = \sum_{\sigma\in S_n}\prod_{i=1}^{n}(1+a_ib_{\sigma(i)})$$

where for each σ, we have

$$\prod_{i=1}^{n}(1+a_ib_{\sigma(i)}) = \sum_{S\subseteq[n]} a^S b^{\sigma(S)}$$

For fixed S with $|S|=k$, as σ ranges over all permutations, $\sigma(S)$ will range over all sets $T\in\binom{[n]}{k}$ and any σ' for which $\sigma'(S)=\sigma(S)$ will give the same term. As there are a total of $k!(n-k)!$ such permutations, we have

$$\operatorname{perm}\left[A\right] = \sum_{k=0}^{n} k!(n-k)!\sum_{S\in\binom{[n]}{k}}\sum_{T\in\binom{[n]}{k}} a^S b^T = \sum_{k=0}^{n} k!(n-k)!e_k(a)e_k(b)$$

as claimed. □

Corollary 2.4.

$$(n!)^2\det\left[A_n\right] = (n^n)\det\left[A_{n-1}\right]\ \operatorname{perm}\left[A\right].$$

Proof. By Lemma 2.1, we have

$$\det\left[A_n\right] = \Delta(a)\Delta(b) \sum_{\substack{S\subseteq\{0,\ldots,n\}\\|S|=n}} s_\lambda(a)s_\lambda(b)\left(\prod_{i\in S}\binom{n}{i}\right).$$

Note that there are n possible subsets in the sum and that each has exactly one element from $\{0,\ldots,n\}$ missing. Letting $R_1=\prod_{i=0}^{n}\binom{n}{i}$ and indexing by the missing element, we can write

$$\det\left[A_n\right] = \Delta(a)\Delta(b)\sum_{t=0}^{n}\frac{R_1}{\binom{n}{t}}s_\lambda(a)s_\lambda(b)$$

where $\lambda_j = (1^t)$. Hence by Lemma 1.3 and then Lemma 2.3, we have

$$\begin{aligned}\det[A_n] &= \Delta(a)\Delta(b)\sum_{t=0}^{n}\frac{R_1}{\binom{n}{t}}e_t(a)e_t(b)\\ &= \frac{R_1}{n!}\Delta(a)\Delta(b)\ \mathrm{perm}[A].\end{aligned}$$

Now if we write

$$R_1 = \prod_{i=0}^{n}\binom{n}{i} = \prod_{i=1}^{n}\frac{n}{i}\binom{n-1}{i-1} = \frac{n^n}{n!}\prod_{i=0}^{n-1}\binom{n-1}{i}$$

then plugging in Corollary 2.2 gives the theorem. □

Theorem 2.5. *Let $X \in \mathbb{F}^{n\times n}$ be any rank 2 matrix and let $X_{n-1}, X_n \in \mathbb{F}^{n\times n}$ be the matrices with*

$$X_{n-1}(i,j) = X(i,j)^{n-1} \qquad \textit{and} \qquad X_n(i,j) = X(i,j)^n$$

Then

$$(n!)^2\det[X_n] = (n^n)\det[X_{n-1}]\ \ \mathrm{perm}[X].$$

Proof. Let $\vec{1} \in \mathbb{F}^n$ be the vector with $\vec{1}(k) = 1$ for all k. Then Corollary 2.4 proves the theorem when

$$X = ab^T + \vec{1}\vec{1}^T.$$

Let $Y = ab^T + cd^T$ for general c, d. Then it is easy to check that the expansion of $\det[Y_n]$ in terms of monomials has the form

$$\det[Y_n] = \sum_{i_1,\dots,i_n,j_1,\dots,j_n} u_{i_1,\dots,i_n,j_1,\dots,j_n}\prod_k a_k^{i_k}c_k^{n-i_k}b_k^{j_k}d_k^{n-j_k} \tag{6}$$

where the $u_{i_1,\dots,i_n,j_1,\dots,j_n}$ are constants. Similarly, $\det[Y_{n-1}]\ \ \mathrm{perm}[Y]$ has an expansion

$$\det[Y_{n-1}]\ \ \mathrm{perm}[Y] = \sum_{i_1,\dots,i_n,j_1,\dots,j_n} v_{i_1,\dots,i_n,j_1,\dots,j_n}\prod_k a_k^{i_k}c_k^{n-i_k}b_k^{j_k}d_k^{n-j_k}. \tag{7}$$

Plugging in $d = \vec{1}$ and $c = \vec{1}$, however, does not cause any of the coefficients to combine. That is,

$$\det[X_n] = \sum_{i_1,\dots,i_n,j_1,\dots,j_n} u_{i_1,\dots,i_n,j_1,\dots,j_n}\prod_k a_k^{i_k}b_k^{j_k}$$

and

$$\det[X_{n-1}]\ \ \mathrm{perm}[X] = \sum_{i_1,\dots,i_n,j_1,\dots,j_n} v_{i_1,\dots,i_n,j_1,\dots,j_n}\prod_k a_k^{i_k}b_k^{j_k}$$

and so by Corollary 2.4, we have

$$(n!)^2 u_{i_1,\dots,i_n,j_1,\dots,j_n} = (n^n)v_{i_1,\dots,i_n,j_1,\dots,j_n}$$

for all indices $i_1,\dots,i_n$ and $j_1,\dots,j_n$. Plugging this into (6) and (7) implies equality for Y. □

References

[1] L. Carlitz and J. Levine. An identity of Cayley. *The American Mathematical Monthly*, 67(6):571–573, 1960.

[2] C. G. J. Jacobi. De functionibus alternantibus earumque divisione per productum e differentiis elementorum conflatum. *Journal für die reine und angewandte Mathematik*, 22:360–371, 1841.

[3] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford university press, 1998.