UNIVERSAL IDENTITIES, II: $\otimes$ AND $\wedge$

KEITH CONRAD

1. Introduction

We will describe how algebraic identities involving operations of multilinear algebra – the tensor product and exterior powers – can be proved by the method of universal identities. Here is an example, showing how the coefficients in the characteristic polynomial of a linear map are related to exterior powers of the linear map.

Theorem 1.1. *Let R be a commutative ring and n be a positive integer. For any $A \in \mathrm{M}_n(R)$, write its characteristic polynomial as*

$$\det(TI_n - A) = T^n + c_1(A)T^{n-1} + \cdots + c_{n-1}(A)T + c_n(A) \in R[T].$$

Then $c_k(A) = (-1)^k \operatorname{Tr}(\wedge^k(A))$.

Note the indexing: the coefficient of T^{n-k} is associated to the kth exterior power of A. In the special cases $k = 1$ and $k = n$ this recovers the more familiar result that $c_1(A) = -\operatorname{Tr}(A)$ and $c_n(A) = (-1)^n \det(A)$. To view the formula for $c_k(A)$ as a universal identity, we need matrix representations for $A \otimes B$ and $\wedge^k(A)$.

Let's recall how tensor products and exterior powers extend from modules to linear maps. Say $A\colon R^n \to R^n$ and $B\colon R^m \to R^m$ are R-linear. Then $A \otimes B$ is the linear operator on $R^n \otimes_R R^m$ which sends $v \otimes w$ to $Av \otimes Bw$, and $\wedge^k(A)$ is the linear operator on $\Lambda^k(R^n)$, for $1 \le k \le n$, which sends any k-fold elementary wedge product $v_1 \wedge \cdots \wedge v_k$ of elements of R^n to the elementary wedge product $A(v_1) \wedge \cdots \wedge A(v_k)$. (We set $\wedge^0(A)$ to be the identity map on $\Lambda^0(R^n) = R$.) Both $R^n \otimes_R R^m$ and $\Lambda^k(R^n)$, for $0 \le k \le n$, admit bases in a definite way from the standard bases $\{e_1, \ldots, e_n\}$ of R^n and $\{f_1, \ldots, f_m\}$ of R^m. The tensor product $R^n \otimes_R R^m$ has the basis

$$e_1 \otimes f_1, \ldots, e_1 \otimes f_m, \ldots, e_n \otimes f_1, \ldots, e_n \otimes f_m,$$

and $\Lambda^k(R^n)$ has the basis $\{e_{i_1} \wedge \cdots \wedge e_{i_k}\}$ with indices in increasing order and arranged lexicographically (for instance, $\Lambda^2(R^3)$ has basis $e_1 \wedge e_2, e_1 \wedge e_3$, and $e_2 \wedge e_3$). The R-linear maps $A \otimes B$ and $\wedge^k(A)$ become concrete matrices relative to these ordered bases. The matrix for $A \otimes B$ in $M_{nm}(R)$ is a partitioned matrix consisting of n^2 different $m \times m$ blocks, where the (i, j) block for $1 \le i, j \le n$ is $a_{ij}B$ as a_{ij} runs over the matrix entries of A in their natural arrangement. (This matrix for $A \otimes B$ is called the "Kronecker product." Look it up on Wikipedia for some examples.) The matrix entries for $\wedge^k(A)$ involve determinants of $k \times k$ submatrices for A but we won't specify precisely where each subdeterminant appears. What matters is that there are definite rules of computation after an ordering of the basis is chosen.

Here are two more theorems about multilinear operations on matrices.

Theorem 1.2. *For $A \in \mathrm{M}_n(R)$ and $B \in \mathrm{M}_m(R)$,*

$$\operatorname{Tr}(A \otimes B) = \operatorname{Tr}(A)\operatorname{Tr}(B) \text{ and } \det(A \otimes B) = \det(A)^m \det(B)^n.$$

Note the exponent on $\det A$ is the size of B and the exponent on $\det B$ is the size of A.

Theorem 1.3 (Sylvester-Franke)**.** *For $A \in \mathrm{M}_n(R)$ and $1 \leq k \leq n$,*

$$\det(\wedge^k(A)) = (\det A)^{\binom{n-1}{k-1}}.$$

To prove these theorems over all commutative rings R, it suffices to treat the case when $R = \mathbf{Z}[X_{11}, \ldots, X_{nn}, Y_{11}, \ldots, Y_{mm}]$, $A = (X_{ij})$, and $B = (Y_{st})$. Then $(X_{ij}) \otimes (Y_{st})$, and $\wedge^k(X_{ij})$ are specific matrices over this ring, and their traces and determinants are in $\mathbf{Z}[X_{11}, \ldots, X_{nn}, Y_{11}, \ldots, Y_{mm}]$. By the method of universal identities, the validity of the theorems follows from the special case of these specific matrices over the specific polynomial ring R, and this special case in turn follows from the special case of complex matrices, where the theorems only need to be checked on an open set of matrices.

2. The proofs

Proof. (of Theorem 1.1) Both $c_k(A)$ and $(-1)^k \operatorname{Tr}(\wedge^k(A))$ are universal polynomials in the matrix entries of A, so it suffices to verify their equality when A is a diagonalizable matrix in $\mathrm{M}_n(\mathbf{C})$. Since the characteristic polynomial of a linear map is independent of the choice of matrix representation, $c_k(A)$ is unchanged if we replace A by a conjugate, and $\operatorname{Tr}(\wedge^k(A))$ is also unchanged by this. Therefore we may take A to be a diagonal matrix, say with diagonal entries $\lambda_1, \ldots, \lambda_n$. Then $Ae_i = \lambda_i e_i$ where $e_1, \ldots, e_n$ is the standard basis of $\mathbf{C}^n$. Since

$$\chi_A(T) = \det(TI_n - A) = \prod_{i=1}^{n}(T - \lambda_i),$$

the coefficient of T^{n-k} is

$$c_k(A) = (-1)^k \sum_{1 \leq i_1 < \cdots < i_k \leq n} \lambda_{i_1} \cdots \lambda_{i_k}.$$

At the same time, $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \leq i_1 < \cdots < i_k \leq n\}$ is an eigenbasis for $\wedge^k(A)$ acting on $\Lambda^k(\mathbf{C}^n)$, where $e_{i_1} \wedge \cdots \wedge e_{i_k}$ has eigenvalue $\lambda_{i_1} \cdots \lambda_{i_k}$ since

$$\wedge^k(A)(e_{i_1} \wedge \cdots \wedge e_{i_k}) = Ae_{i_1} \wedge \cdots \wedge Ae_{i_k} = \lambda_{i_1} e_{i_1} \wedge \cdots \wedge \lambda_{i_k} e_{i_k} = \lambda_{i_1} \cdots \lambda_{i_k}(e_{i_1} \wedge \cdots \wedge e_{i_k}),$$

so

$$\operatorname{Tr}(\wedge^k(A)) = \sum_{1 \leq i_1 < \cdots < i_k \leq n} \lambda_{i_1} \cdots \lambda_{i_k}.$$

Thus $c_k(A) = (-1)^k \operatorname{Tr}(\wedge^k(A))$. □

Proof. (of Theorem 1.2) The identity in matrix pairs $(A, B) \in \mathrm{M}_n(\mathbf{C}) \times \mathrm{M}_n(\mathbf{C})$ will be checked on pairs of diagonalizable matrices, which contains an open set of matrices. Letting A and B be diagonalizable matrices with eigenbases $e_1, \ldots, e_n$ and $f_1, \ldots, f_m$, $Ae_i = \lambda_i e_i$ and $Bf_s = \mu_s f_s$. Then the set $\{e_i \otimes f_s\}$ is a basis of $\mathbf{C}^n \otimes_{\mathbf{C}} \mathbf{C}^m$ and is an eigenbasis for $A \otimes B$ acting on $\mathbf{C}^n \otimes_{\mathbf{C}} \mathbf{C}^m$:

$$(A \otimes B)(e_i \otimes f_s) = (Ae_i) \otimes (Bf_s) = \lambda_i e_i \otimes \mu_s f_s = (\lambda_i \mu_s)(e_i \otimes f_s).$$

The trace and determinant are the sum and product of the eigenevalues (with multiplicity), so

$$\operatorname{Tr}(A \otimes B) = \sum_{i,s} \lambda_i \mu_s = \sum_i \lambda_i \sum_s \mu_s = \operatorname{Tr}(A) \operatorname{Tr}(B)$$

and

$$\begin{aligned}
\det(A \otimes B) &= \prod_{i,s} \lambda_i \mu_s \\
&= \prod_{i=1}^{n} \prod_{s=1}^{m} \lambda_i \mu_s \\
&= \prod_{i=1}^{n} \left(\lambda_i^m \prod_{s=1}^{m} \mu_s \right) \\
&= \prod_{i=1}^{n} (\lambda_i^m (\det B)) \\
&= (\det B)^n \prod_{i=1}^{n} \lambda_i^m \\
&= (\det B)^n (\det A)^m.
\end{aligned}$$

We're done. □

Setting $A = B$, $\mathrm{Tr}(A^{\otimes 2}) = (\mathrm{Tr}\, A)^2$ and $\det(A^{\otimes 2}) = (\det A)^{2n}$. More generally, by induction $\mathrm{Tr}(A^{\otimes k}) = (\mathrm{Tr}\, A)^k$ and $\det(A^{\otimes k}) = (\det A)^{kn^{k-1}}$.

Remark 2.1. If $\chi_A(T) = \prod_i (T - \lambda_i)$ and $\chi_B(T) = \prod_j (T - \mu_j)$, then

$$\chi_{A \otimes B}(T) = \prod_{i,j} (T - \lambda_i \mu_j).$$

Looking at coefficients on both sides recovers Theorem 1.2 for the case of diagonalizable matrices.

Proof. (of Theorem 1.3) We may suppose A is a diagonalizable matrix in $\mathrm{M}_n(\mathbf{C})$ with eigenbasis $e_1, \ldots, e_n$: $Ae_i = \lambda_i e_i$. Then a basis for $\wedge^k(A)$ acting on $\Lambda^k(\mathbf{C}^n)$ is all k-fold elementary wedge products $e_{i_1} \wedge \cdots \wedge e_{i_k}$ $(1 \leq i_1 < \cdots < i_k \leq n)$ and these are eigenvectors for $\wedge^k(A)$:

$$\wedge^k(A)(e_{i_1} \wedge \cdots \wedge e_{i_k}) = \lambda_{i_1} \cdots \lambda_{i_k} (e_{i_1} \wedge \cdots \wedge e_{i_k}). \tag{2.1}$$

Thus

$$\det(\wedge^k(A)) = \prod_{1 \leq i_1 < \cdots < i_k \leq n} \lambda_{i_1} \cdots \lambda_{i_k}.$$

Now we have to unravel this product and figure out how often each λ_i is appearing.

Let $P_{k,n}$ denote the expression on the right side. It is obviously of the form $\lambda_1^{\ell_1} \cdots \lambda_n^{\ell_n}$ for some exponents $\ell_1, \ldots, \ell_n$, and we simply need to show all the exponents equal $\binom{n-1}{k-1}$. This is clear for $k = 1$ and $k = n$, so we may take $2 \leq k \leq n - 1$. The term λ_n appears in $P_{k,n}$ as often as there are $k - 1$ integers $1 \leq i_1 < \cdots < i_{k-1} \leq n - 1$, which is $\binom{n-1}{k-1}$. So

$$\begin{aligned}
P_{k,n} &= \prod_{1 \leq i_1 < \cdots < i_k \leq n-1} (\lambda_{i_1} \cdots \lambda_{i_k}) \cdot \prod_{1 \leq i_1 < \cdots < i_{k-1} \leq n-1} (\lambda_{i_1} \cdots \lambda_{i_{k-1}} \lambda_n) \\
&= P_{k,n-1} P_{k-1,n-1} \lambda_n^{\binom{n-1}{k-1}}.
\end{aligned}$$

By induction on n, assume $P_{k,n-1} = (\lambda_1 \cdots \lambda_{n-1})^{\binom{n-2}{k-1}}$ and $P_{k-1,n-1} = (\lambda_1 \cdots \lambda_{n-1})^{\binom{n-2}{k-2}}$. Since $\binom{n-2}{k-1} + \binom{n-2}{k-2} = \binom{n-1}{k-1}$, we get

$$P_{k,n} = (\lambda_1 \cdots \lambda_{n-1})^{\binom{n-1}{k-1}} \lambda_n^{\binom{n-1}{k-1}} = (\lambda_1 \cdots \lambda_n)^{\binom{n-1}{k-1}},$$

so $\det(\wedge^k(A)) = (\det A)^{\binom{n-1}{k-1}}$. □

We are not discussing symmetric powers, but the methods used on exterior powers can be applied to them too. As an exercise, prove for $A \in \mathrm{M}_n(R)$ and $k \geq 1$ that $\det(\mathrm{Sym}^k(A)) = (\det A)^{\binom{n+k-1}{k-1}}$. For example, $\det(\mathrm{Sym}^2(A)) = (\det A)^{n+1}$.

3. The Consequences

Now we can draw an interesting conclusion about tensor and exterior powers of linear maps.

Corollary 3.1. *Let M be a finite free R-module of rank $n \geq 1$ and $\varphi \colon M \to M$ be linear. Fix a positive integer k. Then φ is an automorphism of M if and only if $\varphi^{\otimes k}$ is an automorphism of $M^{\otimes k}$ and also if and only if $\wedge^k(\varphi)$ is an automorphism of $\Lambda^k(M)$, where $1 \leq k \leq n$ in the case of exterior powers.*

Proof. Since M is free, both $M^{\otimes k}$ and $\Lambda^k(M)$ are free. A linear operator on a finite free R-module is an automorphism if and only if its determinant is in $R^\times$. By Theorems 1.2 and 1.3, $\varphi^{\otimes k}$ and $\wedge^k(\varphi)$ have determinants which are powers of the determinant of φ. An element of R is a unit if and only if some power of it is a unit, so we're done. □

Remark 3.2. That a linear operator on a finite free module is an automorphism if and only if its determinant is a unit can be viewed as the special case $k = n$ of Corollary 3.1 for exterior powers, but we used that special case in the proof.

In the setting of vector spaces, here is an alternate proof of Corollary 3.1. Take V to be a finite-dimensional vector space and $\varphi \colon V \to V$ to be linear. If φ is an automorphism of V then $\varphi^{\otimes k}$ and $\wedge^k(\varphi)$ are automorphisms of $V^{\otimes k}$ and $\Lambda^k(V)$ (their inverses are the kth tensor or exterior power of the inverse of φ). Conversely, suppose φ is not an automorphism of V. Then φ is not one-to-one, so some $v \in V$ with $v \neq 0$ satisfies $\varphi(v) = 0$. Extend v to a basis $v_1, \ldots, v_n$ of V with $v = v_1$. Then the elementary tensor $v_1^{\otimes k}$ is a nonzero element of $V^{\otimes k}$ and, if $k \leq n$, the elementary wedge product $v_1 \wedge v_2 \wedge \cdots \wedge v_k$ is nonzero in $\Lambda^k(V)$. The tensor $v_1^{\otimes k}$ is killed by $\varphi^{\otimes k}$ and this wedge product $v_1 \wedge v_2 \wedge \cdots \wedge v_k$ is killed by $\wedge^k(\varphi)$, so $\varphi^{\otimes k}$ and $\wedge^k(\varphi)$ are not injective and thus are not automorphisms. This proof is not valid on finite free modules over a commutative ring since a nonzero element of a finite free module need not belong to a basis, unlike in the case of vector spaces.

Corollary 3.3. *Let M and N be finite free R-modules of equal rank n and $f \colon M \to N$ be a linear map.*

1. *For each $k \geq 1$, f is an isomorphism if and only if $f^{\otimes k} \colon M^{\otimes k} \to N^{\otimes k}$ is an isomorphism.*
2. *For an integer k with $1 \leq k \leq n$, f is an isomorphism if and only if $\wedge^k(f) \colon \Lambda^k(M) \to \Lambda^k(N)$ is an isomorphism.*
3. *The map f is surjective if and only if $f^{\otimes k}$ is surjective (some $k \geq 1$) or $\wedge^k(f)$ is surjective (some $1 \leq k \leq n$).*

Proof. 1) The direction ($\Rightarrow$) is clear. Conversely, suppose some $f^{\otimes k}$ is an isomorphism. (We just assume this for one k.) We want to show f is an isomorphism.

The modules M and N are isomorphic since they are each isomorphic to R^n. Let $\varphi\colon N \to M$ be an isomorphism and consider the composite map

$$M \xrightarrow{f} N \xrightarrow{\varphi} M.$$

Since φ is an isomorphism, so is $\varphi^{\otimes k}$. Then $(\varphi \circ f)^{\otimes k} = \varphi^{\otimes k} \circ f^{\otimes k}$ is an automorphism of $M^{\otimes k}$. By Corollary 3.1, $\varphi \circ f$ is an automorphism of M, so $f = \varphi^{-1} \circ (\varphi \circ f)$ is an isomorphism.

2) This is similar to part 1.

3) By the theorem of Strooker and Vasconcelos from the first handout on universal identities, a linear map between finite free R-modules of equal rank is surjective if and only if it is an isomorphism. Both f, $f^{\otimes k}$, and $\wedge^k(f)$ are maps between finite free R-modules of equal rank, so by parts 1 and 2 we're done. □

Corollary 3.4. *Let $M \subset N$ be finite free R-modules of equal rank n and $M \neq N$. Let $i\colon M \hookrightarrow N$ be the inclusion map. The maps $i^{\otimes k}\colon M^{\otimes k} \to N^{\otimes k}$ and $\wedge^k(i)\colon \Lambda^k(M) \to \Lambda^k(N)$ are not onto for any $1 \leq k \leq n$.*

Proof. The inclusion is not onto, so we may apply part c of Corollary 3.3. □

Appendix A. Identities with Resultants

For readers who know about resultants of polynomials, we prove some more universal identities.

Theorem A.1. *For $A \in \mathrm{M}_m(R)$ and $B \in \mathrm{M}_n(R)$, let $f(T) = \det(TI_m - A)$ and $g(T) = \det(TI_n - B)$. Then $\det(A \otimes I_n - I_m \otimes B) = \mathrm{Res}(f, g)$ is the resultant of f and g.*

Proof. Both sides are universal polynomials in the entries of A and B. Fix $B \in \mathrm{M}_n(\mathbf{C})$. It suffices to check the identity on diagonal matrices A in $\mathrm{M}_n(\mathbf{C})$. Let $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Then as Kronecker products (the block matrix representation of tensor products of matrices),

$$\begin{aligned} A \otimes I_n - I_m \otimes B &= \begin{pmatrix} \lambda_1 I_n & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & \lambda_n I_n \end{pmatrix} - \begin{pmatrix} B & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & B \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 I_n - B & \cdots & O \\ \vdots & \ddots & \vdots \\ O & \cdots & \lambda_n I_n - B \end{pmatrix}, \end{aligned}$$

which is a block-diagonal matrix. Its determinant is $\prod_{i=1}^{n} \det(\lambda_i I_n - B) = g(\lambda_1) \cdots g(\lambda_n)$, which is $\mathrm{Res}(f, g)$ since f is monic. □

Theorem A.2. *For $A \in \mathrm{M}_n(R)$ and $g(T) \in R[T]$,*

$$\det(g(A)) = \mathrm{Res}(\chi_A(T), g(T)),$$

where Res *is the resultant.*

Proof. Let $A = (X_{ij})$ be a matrix with n^2 indeterminate entries and let $g(T) = Y_dT^d + Y_{d-1}T^{d-1} + \cdots + Y_1T + Y_0$ be a polynomial with indeterminate coefficients. Over the particular ring $\mathbf{Z}[X_{11}, \ldots, X_{nn}, Y_0, \ldots, Y_d]$, Theorem A.2 says

$$\det(g((X_{ij}))) = \mathrm{Res}(\det(TI_n - (X_{ij}), g(T)),$$

which is a polynomial identity because the resultant of two polynomials is a polynomial function of the coefficients of the two polynomials. To prove this identity, it suffices to prove it with $g(T)$ fixed in $\mathbf{C}[T]$ and then letting the matrix $A = (x_{ij})$ run over some set containing an open set in $\mathrm{M}_n(\mathbf{C})$. This will imply the identity is true as a polynomial equality and then it specializes to an identity in all commutative rings.

We may focus on the case when $A \in \mathrm{M}_n(\mathbf{C})$ is diagonalizable. Both sides of the identity are insensitive to replacing A by a conjugate (*e.g.*, on the left side $g(UAU^{-1}) = Ug(A)U^{-1}$ and conjugate matrices have the same determinant, while on the right side $\chi_{UAU^{-1}}(T) = \chi_A(T)$), so we can take A to be diagonal:

$$A = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

Then

$$g(A) = \begin{pmatrix} g(\lambda_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & g(\lambda_n) \end{pmatrix},$$

so

$$\det(g(A)) = g(\lambda_1) \cdots g(\lambda_n).$$

The resultant $\mathrm{Res}(f(T), g(T))$ of two polynomials in given by an integral polynomial in the coefficients of f and g. In the special case when $f(T) = c(T - r_1) \cdots (T - r_n)$, $\mathrm{Res}(f, g) = c^{\deg g} g(r_1) \cdots g(r_n)$.

Since $\chi_A(T) = (T - \lambda_1) \cdots (T - \lambda_n)$ is monic, $\mathrm{Res}(\chi_A, g) = g(\lambda_1) \cdots g(\lambda_n)$, so $\det(g(A)) = \mathrm{Res}(\chi_A, g)$. □

Remark A.3. The proof of Theorem A.2 glided over a technical point: the resultant usually depends on the degrees of the two polynomials involved, so it doesn't always commute with specialization since specialization drops the degree of a polynomial when the leading coefficient is specialized to 0. For example, $\mathrm{Res}(aT + b, cT + d) = ad - bc$ when a and c are nonzero, while $\mathrm{Res}(aT + b, d) = d$ when a and d are nonzero. Note $(ad - bc)|_{c=0} = ad \neq d$ in general! This doesn't bode well for deducing an identity about resultants in all commutative rings by specialization from an identity involving resultants with indeterminate coefficients, as we want to do. However, we are saved by the fact that the characteristic polynomial of (X_{ij}) is a monic polynomial and the resultant of two polynomials doesn't depend on the degrees of the polynomials when one of the polynomials is monic (so formation of such a resultant commutes with specialization). For example, $\mathrm{Res}(T + b, cT + d) = d - bc$ and $\mathrm{Res}(T + b, d) = d$.