17. Vandermonde determinants

1. *Vandermonde determinants*

A rigorous systematic evaluation of Vandermonde determinants (below) of the following identity uses the fact that a polynomial ring over a UFD is again a UFD. A **Vandermonde matrix** is a square matrix of the form in the theorem.

[1.0.1] Theorem:

$$\det \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ x_1^3 & x_2^3 & \dots & x_n^3 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{pmatrix} = (-1)^{n(n-1)/2} \cdot \prod_{i<j} (x_i - x_j)$$

[1.0.2] Remark: The most universal version of the assertion uses indeterminates x_i, and proves an identity in

$$\mathbb{Z}[x_1, \dots, x_n]$$

Proof: First, the idea of the proof. Whatever the determinant may be, it is a polynomial in $x_1, \dots, x_n$. The most universal choice of interpretation of the coefficients is as in $\mathbb{Z}$. If two columns of a matrix are the same, then the determinant is 0. From this we would *want* to conclude that for $i \neq j$ the determinant is *divisible by*[1] $x_i - x_j$ *in the polynomial ring* $\mathbb{Z}[x_1, \dots, x_n]$. If we can conclude that, then, since these polynomials

[1] If one treats the x_i merely as complex numbers, for example, then one *cannot* conclude that the product of the expressions $x_i - x_j$ with $i < j$ divides the determinant. Attempting to evade this problem by declaring the x_i as somehow *variable* complex numbers is an impulse in the right direction, but is made legitimate only by treating *genuine* indeterminates.

are pairwise relatively prime, we can conclude that the determinant is divisible by

$$\prod_{i<j} (x_i - x_j)$$

Considerations of *degree* will show that there is no room for further factors, so, up to a constant, this is the determinant.

To make sense of this line of argument, first observe that a determinant is a polynomial function of its entries. Indeed, the formula is

$$\det M = \sum_p \sigma(p)\, M_{1p(1)} M_{2p(2)} \ldots M_{np(n)}$$

where p runs over permutations of n things and $\sigma(p)$ is the *sign* or *parity* of p, that is, $\sigma(p)$ is $+1$ if p is a product of an *even* number of 2-cycles and is -1 if p is the product of an *odd* number of 2-cycles. Thus, for any $\mathbb{Z}$-algebra homomorphism f to a commutative ring R with identity,

$$f : \mathbb{Z}[x_1, \ldots, x_n] \longrightarrow R$$

we have

$$f(\det V) = \det f(V)$$

where by $f(V)$ we mean application of f entry-wise to the matrix V. Thus, if we can prove an identity in $\mathbb{Z}[x_1, \ldots, x_n]$, then we have a corresponding identity in any ring.

Rather than talking about setting x_j equal to x_i, it is safest to try to see divisibility property as directly as possible. Therefore, we do *not* attempt to use the property that the determinant of a matrix with two equal columns is 0. Rather, we use the property [2] that if an element r of a ring R divides every element of a column (or row) of a square matrix, then it divides the determinant. And we are allowed to add any multiple of one column to another without changing the value of the determinant. Subtracting the j^{th} column from the i^{th} column of our Vandermonde matrix (with $i < j$), we have

$$\det V = \det \begin{pmatrix} \ldots & 1-1 & \ldots & 1 & \ldots \\ \ldots & x_i - x_j & \ldots & x_j & \ldots \\ \ldots & x_i^2 - x_j^2 & \ldots & x_j^2 & \ldots \\ \ldots & x_i^3 - x_j^3 & \ldots & x_j^3 & \ldots \\ & \vdots & & \vdots & \\ \ldots & x_i^{n-1} - x_j^{n-1} & \ldots & x_j^{n-1} & \ldots \end{pmatrix}$$

From the identity

$$x^m - y^m = (x-y)(x^{m-1} + x^{m-2}y + \ldots + y^{m-1})$$

it is clear that $x_i - x_j$ divides all entries of the new i^{th} column. Thus, $x_i - x_j$ divides the determinant. This holds for all $i < j$.

Since these polynomials are linear, they are irreducible in $\mathbb{Z}[x_1, \ldots, x_n]$. Generally, the units in a polynomial ring $R[x_1, \ldots, x_n]$ are the units $R^\times$ in R, so the units in $\mathbb{Z}[x_1, \ldots, x_n]$ are just ± 1. Visibly, the various irreducible $x_i - x_j$ are not *associate*, that is, do not merely differ by units. Therefore, their least common multiple is their product. Since $\mathbb{Z}[x_1, \ldots, x_n]$ is a UFD, this product divides the determinant of the Vandermonde matrix.

To finish the computation, we want to argue that the determinant can have no *further* polynomial factors than the ones we've already determined, so up to a constant (which we'll determine) is equal to the latter

[2] This follows directly from the just-quoted formula for determinants, and also from other descriptions of determinants, but from any viewpoint is still valid for matrices with entries in any commutative ring with identity.

product. [3] To prove this, we need the notion of **total degree**: the total degree of a monomial $x_1^{m_1} \ldots x_n^{m_n}$ is $m_1 + \ldots + m_n$, and the total degree of a polynomial is the maximum of the total degrees of the monomials occurring in it. We grant for the moment the result of the proposition below, that the total degree of a product is the sum of the total degrees of the factors. The total degree of the product is

$$\sum_{1 \le i < j \le n} 1 = \sum_{1 \le i < n} n - i = \frac{1}{2} n(n-1)$$

To determine the total degree of the determinant, invoke the usual formula for the determinant of a matrix M with entries M_{ij}, namely

$$\det M \;=\; \sum_\pi \sigma(\pi) \prod_i M_{i,\pi(i)}$$

where π is summed over permutations of n things, and where $\sigma(\pi)$ is the *sign* of the permutation π. In a Vandermonde matrix all the top row entries have total degree 0, all the second row entries have total degree 1, and so on. Thus, in this permutation-wise sum for a Vandermonde determinant, each summand has total degree

$$0 + 1 + 2 + \ldots + (n-1) = \frac{1}{2} n(n-1)$$

so the total degree of the determinant is the total degree of the product

$$\sum_{1 \le i < j \le n} 1 = \sum_{1 \le i < n} n - i = \frac{1}{2} n(n-1)$$

Thus,

$$\det \begin{pmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ x_1^3 & x_2^3 & \ldots & x_n^3 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{pmatrix} = \text{constant} \cdot \prod_{i<j} (x_i - x_j)$$

Granting this, to determine the constant it suffices to compare a single monomial in both expressions. For example, compare the coefficients of

$$x_1^{n-1} x_2^{n-2} x_3^{n-3} \ldots x_{n-1}^1 x_n^0$$

In the product, the only way x_1^{n-1} appears is by choosing the x_1s in the linear factors $x_1 - x_j$ with $1 < j$. After this, the only way to get x_2^{n-2} is by choosing all the x_2s in the linear factors $x_2 - x_j$ with $2 < j$. Thus, this monomial has coefficient $+1$ in the product.

In the determinant, the only way to obtain this monomial is as the product of entries from lower left to upper right. The indices of these entries are $(n, 1)$, $(n-1, 2), \ldots, (2, n-1), (1, n)$. Thus, the coefficient of this monomial is $(-1)^\ell$ where ℓ is the number of 2-cycles necessary to obtain the permutation p such that

$$p(i) = n + 1 - i$$

Thus, for n even there are $n/2$ two-cycles, and for n odd $(n-1)/2$ two-cycles. For a closed form, as these expressions will appear only as exponents of -1, we only care about values modulo 2. Because of the division by 2, we only care about n modulo 4. Thus, we have values

$$\begin{cases} n/2 & = & 0 \bmod 2 & (\text{for } n = 0 \bmod 4) \\ (n-1)/2 & = & 0 \bmod 2 & (\text{for } n = 1 \bmod 4) \\ n/2 & = & 1 \bmod 2 & (\text{for } n = 3 \bmod 4) \\ (n-1)/2 & = & 1 \bmod 2 & (\text{for } n = 1 \bmod 4) \end{cases}$$

[3] This is more straightforward than setting up the right viewpoint for the first part of the argument.

After some experimentation, we find a closed expression

$$n(n-1)/2 \bmod 2$$

Thus, the leading constant is

$$(-1)^{n(n-1)/2}$$

in the expression for the Vandermonde determinant. ///

Verify the property of total degree:

[1.0.3] **Lemma:** Let $f(x_1, \ldots, x_n)$ and $g(x_1, \ldots, x_n)$ be polynomials in $k[x_1, \ldots, x_n]$ where k is a field. Then the total degree of the product is the sum of the total degrees.

Proof: It is clear that the total degree of the product is less than or equal the sum of the total degrees.

Let $x_1^{e_1} \ldots x_n^{e_n}$ and $x_1^{f_1} \ldots x_n^{f_n}$ be two monomials of highest total degrees $s = e_1 + \ldots + e_n$ and $t = f_1 + \ldots + f_n$ occurring with non-zero coefficients in f and g, respectively. Assume without loss of generality that the exponents e_1 and f_1 of x_1 in the two expressions are the largest among all monomials of total degrees s, t in f and g, respectively. Similarly, assume without loss of generality that the exponents e_2 and f_2 of x_2 in the two expressions are the largest among all monomials of total degrees s, t in f and g, respectively, of degrees e_1 and f_1 in x_1. Continuing similarly, we claim that the coefficient of the monomial

$$M = x^{e_1+f_1} \ldots x_n^{e_n+f_n}$$

is simply the product of the coefficients of $x_1^{e_1} \ldots x_n^{e_n}$ and $x_1^{f_1} \ldots x_n^{f_n}$, so non-zero. Let $x_1^{u_1} \ldots x_n^{u_n}$ and $x_1^{v_1} \ldots x_n^{v_n}$ be two other monomials occurring in f and g such that for all indices i we have $u_i + v_i = e_i + f_i$. By the maximality assumption on e_1 and f_1, we have $e_1 \geq u_1$ and $f_1 \geq v_1$, so the only way that the necessary power of x_1 can be achieved is that $e_1 = u_1$ and $f_1 = v_1$. Among exponents with these maximal exponents of x_1, e_2 and f_2 are maximal, so $e_2 \geq u_2$ and $f_2 \geq v_2$, and again it must be that $e_2 = u_2$ and $f_2 = v_2$ to obtain the exponent of x_2. Inductively, $u_i = e_i$ and $v_i = f_i$ for all indices. That is, the only terms in f and g contributing to the coefficient of the monomial M in $f \cdot g$ are monomials $x_1^{e_1} \ldots x_n^{e_n}$ and $x_1^{f_1} \ldots x_n^{f_n}$. Thus, the coefficient of M is non-zero, and the total degree is as claimed. ///

---

2. *Worked examples*

[17.1] Show that a *finite* integral domain is necessarily a *field*.

Let R be the integral domain. The integral domain property can be immediately paraphrased as that for $0 \neq x \in R$ the map $y \longrightarrow xy$ has trivial kernel (as R-module map of R to itself, for example). Thus, it is injective. Since R is a finite set, an injective map of it to itself is a bijection. Thus, there is $y \in R$ such that $xy = 1$, proving that x is invertible. ///

[17.2] Let $P(x) = x^3 + ax + b \in k[x]$. Suppose that $P(x)$ factors into linear polynomials $P(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$. Give a polynomial condition on a, b for the α_i to be distinct.

(One might try to do this as a symmetric function computation, but it's a bit tedious.)

If $P(x) = x^3 + ax + b$ has a repeated factor, then it has a common factor with its derivative $P'(x) = 3x^2 + a$.

If the characteristic of the field is 3, then the derivative is the constant a. Thus, if $a \neq 0$, $\gcd(P, P') = a \in k^\times$ is never 0. If $a = 0$, then the derivative is 0, and all the α_i are the same.

Now suppose the characteristic is not 3. In effect applying the Euclidean algorithm to P and P',

$$\left(x^3 + ax + b\right) - \frac{x}{3} \cdot \left(3x^2 + a\right) = ax + b - \frac{x}{3} \cdot a = \frac{2}{3}ax + b$$

If $a = 0$ then the Euclidean algorithm has already terminated, and the condition for distinct roots or factors is $b \neq 0$. Also, possibly surprisingly, at this point we need to consider the possibility that the characteristic is 2. If so, then the remainder is b, so if $b \neq 0$ the roots are always distinct, and if $b = 0$

Now suppose that $a \neq 0$, and that the characteristic is not 2. Then we can divide by $2a$. Continue the algorithm

$$(3x^2 + a) - \frac{9x}{2a} \cdot \left(\frac{2}{3}ax + b\right) = a + \frac{27b^2}{4a^2}$$

Since $4a^2 \neq 0$, the condition that P have no repeated factor is

$$4a^3 + 27b^2 \neq 0$$

[17.3] The first three **elementary symmetric functions** in indeterminates $x_1, \ldots, x_n$ are

$$\sigma_1 = \sigma_1(x_1, \ldots, x_n) = x_1 + x_2 + \ldots + x_n = \sum_i x_i$$

$$\sigma_2 = \sigma_2(x_1, \ldots, x_n) = \sum_{i<j} x_i x_j$$

$$\sigma_3 = \sigma_3(x_1, \ldots, x_n) = \sum_{i<j<\ell} x_i x_j x_\ell$$

Express $x_1^3 + x_2^3 + \ldots + x_n^3$ in terms of $\sigma_1, \sigma_2, \sigma_3$.

Execute the algorithm given in the proof of the theorem. Thus, since the degree is 3, if we can derive the right formula for just 3 indeterminates, the same expression in terms of elementary symmetric polynomials will hold generally. Thus, consider $x^3 + y^3 + z^3$. To approach this we first take $y = 0$ and $z = 0$, and consider x^3. This is $s_1(x)^3 = x^3$. Thus, we next consider

$$(x^3 + y^3) - s_1(x, y)^3 = 3x^2y + 3xy^2$$

As the algorithm assures, this is divisible by $s_2(x, y) = xy$. Indeed,

$$(x^3 + y^3) - s_1(x, y)^3 = (3x + 3y)s_2(x, y) = 3s_1(x, y)\, s_2(x, y)$$

Then consider

$$(x^3 + y^3 + z^3) - \left(s_1(x, y, z)^3 - 3\, s_2(x, y, z)\, s_1(x, y, z)\right) = 3xyz = 3s_3(x, y, z)$$

Thus, again, since the degree is 3, this formula for 3 variables gives the general one:

$$x_1^3 + \ldots + x_n^3 = s_1^3 - 3s_1s_2 + 3s_3$$

where $s_i = s_i(x_1, \ldots, x_n)$.

[17.4] Express $\sum_{i \neq j} x_i^2 x_j$ as a polynomial in the elementary symmetric functions of $x_1, \ldots, x_n$.

We could (as in the previous problem) execute the algorithm that proves the theorem asserting that every symmetric (that is, S_n-invariant) polynomial in $x_1, \ldots, x_n$ is a polynomial in the elementary symmetric functions.

But, also, sometimes *ad hoc* manipulations can yield shortcuts, depending on the context. Here,

$$\sum_{i \neq j} x_i^2 x_j = \sum_{i,j} x_i^2 x_j - \sum_{i=j} x_i^2 x_j = \Big(\sum_i x_i^2\Big)\Big(\sum_j x_j\Big) - \sum_i x_i^3$$

An easier version of the previous exercise gives

$$\sum_i x_i^2 = s_1^2 - 2s_2$$

and the previous exercise itself gave

$$\sum_i x_i^3 = s_1^3 - 3s_1s_2 + 3s_3$$

Thus,

$$\sum_{i\neq j} x_i^2 x_j = (s_1^2 - 2s_2)\, s_1 - \left(s_1^3 - 3s_1s_2 + 3s_3\right) = s_1^3 - 2s_1s_2 - s_1^3 + 3s_1s_2 - 3s_3 = s_1s_2 - 3s_3$$

[17.5] Suppose the characteristic of the field k does not divide n. Let $\ell > 2$. Show that

$$P(x_1, \ldots, x_n) = x_1^n + \ldots + x_\ell^n$$

is irreducible in $k[x_1, \ldots, x_\ell]$.

First, treating the case $\ell = 2$, we claim that $x^n + y^n$ is not a unit and has no repeated factors in $k(y)[x]$. (We take the field of rational functions in y so that the resulting polynomial ring in a single variable is Euclidean, and, thus, so that we understand the behavior of its irreducibles.) Indeed, if we start executing the Euclidean algorithm on $x^n + y^n$ and its derivative nx^{n-1} in x, we have

$$(x^n + y^n) - \frac{x}{n}(nx^{n-1}) = y^n$$

Note that n is invertible in k by the characteristic hypothesis. Since y is invertible (being non-zero) in $k(y)$, this says that the *gcd* of the polynomial in x and its derivative is 1, so there is no repeated factor. And the degree in x is positive, so $x^n + y^n$ has *some* irreducible factor (due to the unique factorization in $k(y)[x]$, or, really, due indirectly to its Noetherian-ness).

Thus, our induction (on n) hypothesis is that $x_2^n + x_3^n + \ldots + x_\ell^n$ is a non-unit in $k[x_2, x_3, \ldots, x_n]$ and has no repeated factors. That is, it is divisible by some irreducible p in $k[x_2, x_3, \ldots, x_n]$. Then in

$$k[x_2, x_3, \ldots, x_n][x_1] \approx k[x_1, x_2, x_3, \ldots, x_n]$$

Eisenstein's criterion applied to $x_1^n + \ldots$ as a polynomial in x_1 with coefficients in $k[x_2, x_3, \ldots, x_n]$ and using the irreducible p yields the irreducibility.

[17.6] Find the determinant of the **circulant** matrix

$$\begin{pmatrix} x_1 & x_2 & \ldots & x_{n-2} & x_{n-1} & x_n \\ x_n & x_1 & x_2 & \ldots & x_{n-2} & x_{n-1} \\ x_{n-1} & x_n & x_1 & x_2 & \ldots & x_{n-2} \\ \vdots & & & \ddots & & \vdots \\ x_3 & & & & x_1 & x_2 \\ x_2 & x_3 & \ldots & & x_n & x_1 \end{pmatrix}$$

(*Hint:* Let ζ be an n^{th} root of 1. If $x_{i+1} = \zeta \cdot x_i$ for all indices $i < n$, then the $(j+1)^{th}$ row is ζ times the j^{th}, and the determinant is 0.)

Let C_{ij} be the ij^{th} entry of the circulant matrix C. The expression for the determinant

$$\det C = \sum_{p\in S_n} \sigma(p)\, C_{1,p(1)} \ldots C_{n,p(n)}$$

where $\sigma(p)$ is the sign of p shows that the determinant is a polynomial in the entries C_{ij} with integer coefficients. This is the most universal viewpoint that could be taken. However, with some hindsight, some intermediate manipulations suggest or require enlarging the 'constants' to include n^{th} roots of unity ω. Since we do not know that $\mathbb{Z}[\omega]$ is a UFD (and, indeed, it is not, in general), we must adapt. A reasonable adaptation is to work over $\mathbb{Q}(\omega)$. Thus, we will prove an identity in $\mathbb{Q}(\omega)[x_1, \ldots, x_n]$.

Add ω^{i-1} times the i^{th} row to the first row, for $i \geq 2$. The new first row has entries, from left to right,

$$x_1 + \omega x_2 + \omega^2 x_3 + \ldots + \omega^{n-1} x_n$$

$$x_2 + \omega x_3 + \omega^2 x_4 + \ldots + \omega^{n-1} x_{n-1}$$

$$x_3 + \omega x_4 + \omega^2 x_5 + \ldots + \omega^{n-1} x_{n-2}$$

$$x_4 + \omega x_5 + \omega^2 x_6 + \ldots + \omega^{n-1} x_{n-3}$$

$$\ldots$$

$$x_2 + \omega x_3 + \omega^2 x_4 + \ldots + \omega^{n-1} x_1$$

The t^{th} of these is

$$\omega^{-t} \cdot (x_1 + \omega x_2 + \omega^2 x_3 + \ldots + \omega^{n-1} x_n)$$

since $\omega^n = 1$. Thus, in the ring $\mathbb{Q}(\omega)[x_1, \ldots, x_n]$,

$$x_1 + \omega x_2 + \omega^2 x_3 + \ldots + \omega^{n-1} x_n)$$

divides this new top row. Therefore, from the explicit formula, for example, this quantity divides the determinant.

Since the characteristic is 0, the n roots of $x^n - 1 = 0$ are distinct (for example, by the usual computation of *gcd* of $x^n - 1$ with its derivative). Thus, there are n superficially-different linear expressions which divide $\det C$. Since the expressions are linear, they are *irreducible* elements. If we prove that they are *non-associate* (do not differ merely by units), then their product must divide $\det C$. Indeed, viewing these linear expressions in the larger ring

$$\mathbb{Q}(\omega)(x_2, \ldots, x_n)[x_1]$$

we see that they are distinct linear monic polynomials in x_1, so are non-associate.

Thus, for some $c \in \mathbb{Q}(\omega)$,

$$\det C = c \cdot \prod_{1 \leq \ell \leq n} \left(x_1 + \omega^\ell x_2 + \omega^{2\ell} x_3 + \omega^{3\ell} x_4 + \ldots + \omega^{(n-1)\ell} x_n \right)$$

Looking at the coefficient of x_1^n on both sides, we see that $c = 1$.

(One might also observe that the product, when expanded, will have coefficients in $\mathbb{Z}$.)

Exercises

17.[2.0.1] A k-linear *derivation* D on a commutative k-algebra A, where k is a field, is a k-linear map $D : A \longrightarrow A$ satisfying *Leibniz' identity*

$$D(ab) = (Da) \cdot b + a \cdot (Db)$$

Given a polynomial $P(x)$, show that there is a unique k-linear derivation D on the polynomial ring $k[x]$ sending x to $P(x)$.

17.[2.0.2] Let A be a commutative k-algebra which is an integral domain, with field of fractions K. Let D be a k-linear derivation on A. Show that there is a unique extension of D to a k-linear derivation on K, and that this extension necessarily satisfies the quotient rule.

17.[2.0.3] Let $f(x_1, \ldots, x_n)$ be a homogeneous polynomial of *total degree* n, with coefficients in a field k. Let $\partial/\partial x_i$ be partial differentiation with respect to x_i. Prove *Euler's identity*, that

$$\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = n \cdot f$$

17.[2.0.4] Let α be algebraic over a field k. Show that any k-linear derivation D on $k(\alpha)$ necessarily gives $D\alpha = 0$.