THE RESULTANT

1. Newton's identities

The monic polynomial p with roots $r_1, \dots, r_n$ expands as

$$p(T) = \prod_{i=1}^{n}(T - r_i) = \sum_{j \in \mathbb{Z}} (-1)^j \sigma_j T^{n-j} \in \mathbb{C}(\sigma_1, \dots, \sigma_n)[T]$$

whose coefficients are (up to sign) the *elementary symmetric functions* of the roots $r_1, \dots, r_n$,

$$\sigma_j = \sigma_j(r_1, \dots, r_n) = \begin{cases} \sum_{1 \le i_1 < \cdots < i_j \le n} \prod_{k=1}^{j} r_{i_k} & \text{for } j \ge 0 \\ 0 & \text{for } j < 0. \end{cases}$$

In less dense notation,

$$\begin{aligned}
\sigma_1 &= r_1 + \cdots + r_n, \\
\sigma_2 &= r_1 r_2 + r_1 r_3 \cdots + r_{n-1} r_n \quad \text{(the sum of all distinct pairwise products)}, \\
\sigma_3 &= \text{the sum of all distinct triple products}, \\
&\ \vdots \\
\sigma_n &= r_1 \cdots r_n \quad \text{(the only distinct } n\text{-fold product)}.
\end{aligned}$$

Note that $\sigma_0 = 1$ and $\sigma_j = 0$ for $j > n$. The product form of p shows that the σ_j are invariant under all permutations of $r_1, \dots, r_n$.

The *power sums* of $r_1, \dots, r_n$ are

$$s_j = s_j(r_1, \dots, r_n) = \begin{cases} \sum_{i=1}^{n} r_i^j & \text{for } j \ge 0 \\ 0 & \text{for } j < 0 \end{cases}$$

including $s_0 = n$. That is,

$$\begin{aligned}
s_1 &= r_1 + \cdots + r_n \, (= \sigma_1), \\
s_2 &= r_1^2 + r_2^2 + \cdots + r_n^2, \\
&\ \vdots \\
s_n &= r_1^n + \cdots + r_n^n,
\end{aligned}$$

and the s_j for $j > n$ do *not* vanish. Like the elementary symmetric functions σ_j, the power sums s_j are invariant under all permutations of $r_1, \dots, r_n$. We want to relate the s_j to the σ_j.

Start from the general polynomial,

$$p(T) = \prod_{i=1}^{n}(T - r_i) = \sum_{j \in \mathbb{Z}} (-1)^j \sigma_j T^{n-j}.$$

Certainly

$$p'(T) = \sum_{j\in\mathbb{Z}} (-1)^j \sigma_j (n-j) T^{n-j-1}.$$

But also, the logarithmic derivative and geometric series formulas,

$$\frac{p'(T)}{p(T)} = \sum_{i=1}^{n} \frac{1}{T-r_i} \qquad \text{and} \qquad \frac{1}{T-r} = \sum_{k=0}^{\infty} \frac{r^k}{T^{k+1}},$$

give

$$\begin{aligned} p'(T) = p(T)\cdot\frac{p'(T)}{p(T)} &= p(T)\sum_{i=1}^{n}\sum_{k=0}^{\infty}\frac{r_i^k}{T^{k+1}} = p(T)\sum_{k\in\mathbb{Z}}\frac{s_k}{T^{k+1}} \\ &= \sum_{k,l\in\mathbb{Z}} (-1)^l \sigma_l s_k T^{n-k-l-1} \\ &= \sum_{j\in\mathbb{Z}} \left[\sum_{l\in\mathbb{Z}} (-1)^l \sigma_l s_{j-l}\right] T^{n-j-1} \quad \text{(letting } j = k+l\text{).} \end{aligned}$$

Equate the coefficients of the two expressions for p' to get the formula

$$\sum_{l=0}^{j-1} (-1)^l \sigma_l s_{j-l} + (-1)^j \sigma_j n = (-1)^j \sigma_j (n-j).$$

Newton's identities follow,

$$\sum_{l=0}^{j-1} (-1)^l \sigma_l s_{j-l} + (-1)^j \sigma_j j = 0 \quad \text{for all } j.$$

Explicitly, Newton's identities are

$$\begin{aligned} &s_1 - \sigma_1 = 0 \\ &s_2 - s_1\sigma_1 + 2\sigma_2 = 0 \\ &s_3 - s_2\sigma_1 + s_1\sigma_2 - 3\sigma_3 = 0 \\ &s_4 - s_3\sigma_1 + s_2\sigma_2 - s_1\sigma_3 + 4\sigma_4 = 0 \\ &\text{and so on.} \end{aligned}$$

These show (exercise) that for any $j \in \{1, \dots, n\}$, the power sums s_1 through s_j are polynomials (with constant terms zero) in the elementary symmetric functions σ_1 through σ_j, and—since we are in characteristic zero—that the elementary symmetric functions σ_1 through σ_j are polynomials (with constant terms zero) in the power sums s_1 through s_j. Consequently,

Proposition 1.1. *Consider a polynomial*

$$p(T) = T^n + a_1 T^{n-1} + \cdots + a_n.$$

Its first j coefficients $a_1, \dots, a_j$ are zero exactly when the first j power sums of its roots vanish.

Exercises:

- Express s_j in terms of $\sigma_1, \dots, \sigma_j$ for $j = 1, 2, 3$, and conversely.
- Write some of Newton's identities when $j > n$; what is the pattern?

- True or false: the second coefficient a_2 of the polynomial $p(T) = T^n + a_1T^{n-1} + \cdots + a_n$ is zero exactly when the second power sum of its roots vanishes.
- Show that for any $j \in \{1, \ldots, n\}$, the power sums $s_1, \ldots, s_j$ are polynomials (with constant term zero) in the elementary symmetric functions $\sigma_1, \ldots, \sigma_j$, and conversely. (The converse fails in nonzero characteristic; for example, consider $p(T) = T^2 + 1$ in characteristic 2.)
- Establish the formula for the *Vandermonde determinant*,
$$\begin{vmatrix} 1 & r_1 & r_1^2 & \cdots & r_1^{n-1} \\ 1 & r_2 & r_2^2 & \cdots & r_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & r_n & r_n^2 & \cdots & r_n^{n-1} \end{vmatrix} = \prod_{i<j}(r_j - r_i).$$
(Replace the last column by $(p(r_1), \ldots, p(r_n))$, $p(T) = \prod_{i=1}^{n-1}(T - r_i)$.) Left-multiply the Vandermonde matrix by its transpose and take determinants to obtain
$$\begin{vmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-2} \end{vmatrix} = \Delta(r_1, \ldots, r_n),$$
where
$$\Delta(r_1, \ldots, r_n) = \prod_{i<j}(r_i - r_j)^2$$
is the *discriminant* of p. This expresses the discriminant in terms of the elementary symmetric functions $\sigma_1, \ldots, \sigma_n$ since Newton's identities give expressions for the power sums s_j in terms of the σ_j. A formula for the discriminant that doesn't require Newton's identities will be developed in the next section.
- Show that the n-by-n Jacobian matrix of the elementary symmetric functions has the same determinant as the Vandermonde matrix,
$$\det[D_j\sigma_i(r_1, \ldots, r_n)] = \prod_{i<j}(r_j - r_i).$$
Show also that $D_j\sigma_i = \sigma_{i-1}(r_1, \ldots, \overline{r}_j, \ldots, r_n)$, where the overbar means the variable is omitted.

2. Resultants

Given polynomials $p(T)$ and $q(T)$, we can determine whether they have a root in common without actually finding their roots.

Let m and n be nonnegative integers, let $a_0, \ldots, a_m$, $b_0, \ldots, b_n$ be symbols (possibly elements of the base field $\mathbb{C}$) with $a_0 \neq 0$ and $b_0 \neq 0$, and let $\mathbf{k} = \mathbb{C}(a_0, \ldots, a_m, b_0, \ldots, b_n)$. The polynomials
$$p(T) = \sum_{i=0}^{m} a_iT^{m-i} \quad \text{and} \quad q(T) = \sum_{i=0}^{n} b_iT^{n-i}$$
in $\mathbf{k}[T]$ are utterly general when the a_i's and the b_i's form an algebraically independent set, or conversely they can be explicit polynomials when all the coefficients

are in $\mathbb{C}$. The polynomials p and q share a nonconstant factor in $\mathbf{k}[T]$ if and only if there exist nonzero polynomials in $\mathbf{k}[T]$

$$P(T) = \sum_{i=0}^{n-1} c_i T^{n-1-i}, \ \deg(P) < n \quad \text{and} \quad Q(T) = \sum_{i=0}^{m-1} d_i T^{m-1-i}, \ \deg(Q) < m$$

such that

$$pP = qQ.$$

Such P and Q exist if and only if the system

$$vM = 0$$

of $m+n$ linear equations over $\mathbf{k}$ in $m+n$ unknowns has a nonzero solution v, where

$$v = [c_0, c_1, \dots, c_{n-1}, -d_0, -d_1, \dots, -d_{m-1}] \in \mathbf{k}^{m+n},$$

and M is the *Sylvester matrix*,

$$M = \begin{bmatrix} a_0 & a_1 & \cdots & \cdots & a_m & & \\ & \ddots & \ddots & & & \ddots & \\ & & a_0 & a_1 & \cdots & \cdots & a_m \\ b_0 & b_1 & \cdots & b_n & & & \\ & b_0 & b_1 & \cdots & b_n & & \\ & & \ddots & \ddots & & \ddots & \\ & & & b_0 & b_1 & \cdots & b_n \end{bmatrix} \in \mathbf{k}^{(m+n)\times(m+n)},$$

with n staggered rows of a_i's, m staggered rows of b_j's, all other entries 0. Such a nonzero solution exists in turn if and only if $\det M = 0$. This determinant is called the *resultant* of p and q, and it is written $R(p,q)$,

$$R(p,q) = \det(M) \in \mathbf{k} = \mathbb{C}[a_0, \dots, a_m, b_0, \dots, b_n].$$

The condition that $p(T)$ and $q(T)$ share a factor in $\mathbf{k}[T]$ is equivalent to their sharing a root in the splitting field over $\mathbf{k}$ of pq. Thus the result is

Theorem 2.1. *The polynomials $p(T)$ and $q(T)$ in $\mathbf{k}[T]$ share a nonconstant factor in $\mathbf{k}[T]$, or equivalently, share a root in the splitting field over $\mathbf{k}$ of their product, if and only if $R(p,q) = 0$.*

Again, this is handy since $R(p,q)$ is a calculable expression in the coefficients of p and q that makes no reference to the roots. When the coefficients of p and q are algebraically independent, $R(p,q)$ is a master formula that applies to all polynomials of degrees m and n. At the other extreme, when the coefficients are specific values in $\mathbb{C}$, the resultant $R(p,q)$ is a complex number that is zero or nonzero depending on whether the particular polynomials p and q share a factor.

Taking the resultant of p and q to check whether they share a root can also be viewed as eliminating the variable T from the pair of equations $p(T) = 0$, $q(T) = 0$, leaving one equation $R(p,q) = 0$ in the remaining variables $a_0, \dots, a_m, b_0, \dots, b_n$.

In principle, evaluating $R(p,q) = \det M$ can be carried out via row and column operations. In practice, evaluating a large determinant by hand is an error-prone process. The next theorem will supply as a corollary a more efficient method to compute $R(p,q)$. In any case, since any worthwhile computer symbolic algebra package is equipped with a resultant function, nontrivial resultants can often be found by machine.

In their splitting field over $\mathbf{k}$, the polynomials p and q factor as

$$p(T) = a_0 \prod_{i=1}^{m} (T - r_i), \qquad q(T) = b_0 \prod_{j=1}^{n} (T - s_j).$$

To express the resultant $R(p, q)$ explicitly in terms of the roots of p and q, introduce the quantity

$$\tilde{R}(p, q) = a_0^n b_0^m \prod_{i=1}^{m} \prod_{j=1}^{n} (r_i - s_j).$$

This polynomial vanishes if and only if p and q share a root, so it divides $R(p, q)$. Note that $\tilde{R}(p, q)$ is homogeneous of degree mn in the r_i and s_j. On the other hand, each coefficient $a_i = a_0(-1)^i \sigma_i(r_1, \ldots, r_m)$ of p has homogeneous degree i in $r_1, \ldots, r_m$, and similarly for each b_j and $s_1, \ldots, s_n$. Thus in the Sylvester matrix the (i, j)th entry has degree

$$\begin{cases} j - i \text{ in the } r_i & \text{if } 1 \leq i \leq n,\ i \leq j \leq i + m, \\ j - i + n \text{ in the } s_j & \text{if } n + 1 \leq i \leq n + m,\ i - n \leq j \leq i. \end{cases}$$

It quickly follows that any nonzero term in the determinant $R(p, q)$ has degree mn in the r_i and the s_j, and so $\tilde{R}(p, q)$ and $R(p, q)$ agree up to multiplicative constant. Matching coefficients of $(s_1 \cdots s_n)^m$ shows that the constant is 1. This proves

Theorem 2.2. *The resultant of the polynomials*

$$p = \sum_{i=0}^{m} a_i T^{m-i} = a_0 \prod_{i=1}^{m} (T - r_i) \quad \text{and} \quad q = \sum_{j=0}^{n} b_j T^{n-j} = b_0 \prod_{j=1}^{n} (T - s_j)$$

is given by the formulas

$$R(p, q) = a_0^n b_0^m \prod_{i=1}^{m} \prod_{j=1}^{n} (r_i - s_j) = a_0^n \prod_{i=1}^{m} q(r_i) = (-1)^{mn} b_0^m \prod_{j=1}^{n} p(s_j).$$

A special case of this theorem gives the efficient formula for the discriminant promised earlier. See exercise 4.

Computing resultants can now be carried out via a Euclidean algorithm procedure: repeatedly do polynomial division with remainder and apply formula (4) in

Corollary 2.3. *The following formulas hold:*

(1) $R(q, p) = (-1)^{mn} R(p, q)$.
(2) $R(p\tilde{p}, q) = R(p, q) R(\tilde{p}, q)$ *and* $R(p, q\tilde{q}) = R(p, q) R(p, \tilde{q})$.
(3) $R(a_0, q) = a_0^n$ *and* $R(a_0 T + a_1, q) = a_0^n q(-a_1/a_0)$.
(4) *If* $q = Qp + \tilde{q}$ *with* $\deg(\tilde{q}) < \deg(p)$ *then*

$$R(p, q) = a_0^{\deg(q) - \deg(\tilde{q})} R(p, \tilde{q}).$$

Exercise 5 asks for the proofs.

Exercises:

- Show that p and q share a nonconstant factor in $\mathbf{k}[T]$ if and only if there exist nonzero polynomials P of degree less than n and Q of degree less than m in $\mathbf{k}[T]$ such that $pP = qQ$.
- Write out the matrix M for various small values of m and n, and compute the corresponding resultants.

- Fill in the details of the proof of Theorem 2.2.
- (a) Use Theorem 2.2 to show that if p is monic, so that consequently $p' = \sum_{i=1}^{n} \prod_{j \neq i}(T - r_j)$, then

$$R(p, p') = (-1)^{n(n-1)/2} \Delta(p).$$

This formula gives the relation between the resultant and the discriminant.

(b) Use part (a) to recompute the discriminants of $p = T^2 + bT + c$ and of $p = T^3 + bT + c$.

- (a) Prove the formulas in Corollary 2.3.

(b) Let $p = T^n + bT + c$. Compute $\Delta(p) = (-1)^{n(n-1)/2} R(p, p')$ by using the corollary. (Do a polynomial division and apply the second formula in Corollary 2.3. The answer is

$$(-1)^{(n-1)(n-2)/2}(n-1)^{n-1}b^n + (-1)^{n(n-1)/2}n^n c^{n-1}.$$

Note that since n is a general symbol here, evaluating $R(p, p')$ as a determinant is much more awkward than this method.