Symmetric Polynomials: Polys

Jonathan L.F. King
University of Florida, Gainesville FL 32611-2082, USA
squash@ufl.edu
Webpage http://squash.1gainesville.com/
14 October, 2017 (at *11:56*)

ABSTRACT: Proves the Fund. Thm of Symmetric Polynomials, that each sympoly is determined by the *elementary* symmetric polynomials; the *esps*. Indeed, each N-variable sympoly is a polynomial in the N-variable esps.

Along the way, this note provides an introduction to multinomial coefficients.

(See "~/Elisp/SymPoly/symmetricpoly.el for code.)

Review of multinomial coefficients. For a natnum n, use "$n!$" to mean "*n **factorial***"; the product of all posints $\leqslant n$. So $3! = 3 \cdot 2 \cdot 1 = 6$ and $5! = 120$. Also $0! = 1$ and $1! = 1$.

The ***binomial coefficient*** $\binom{7}{3}$, read "7 choose 3", means *the number of ways of choosing* 3 *objects from* 7 *distinguishable objects*. If we think of putting these objects in our left pocket, and putting the remaining 4 objects in our right pocket, then we write the coefficient as $\binom{7}{3,4}$. [Read as "7 *choose* 3-*comma*-4."] Note that $\binom{7}{0} = \binom{7}{0,7} = 1$. Observe that $\binom{N+1}{k+1} = \binom{N}{k} + \binom{N}{k+1}$. Finally, the Binomial theorem says

B1: $$[x+y]^N = \sum_{j+k=N} \binom{N}{j,k} \cdot x^j y^k ,$$

where (j,k) ranges over all *ordered* pairs of natural numbers with sum N.

In general, for natnums $N = k_1 + \ldots + k_P$, the ***multinomial coefficient*** $\binom{N}{k_1,k_2,\ldots,k_P}$ is the number of ways of partitioning N objects, by putting k_1 objects in pocket-one, k_2 objects in pocket-two, …putting k_j objects in the j^{th} pocket. Easily

B2: $$\binom{N}{k_1,k_2,\ldots,k_P} = \frac{N!}{k_1! \cdot k_2! \cdot \ldots \cdot k_P!} .$$

And $[x_1 + \cdots + x_P]^N$ indeed equals the sum of terms

$$\binom{N}{k_1,\ldots,k_P} \cdot {x_1}^{k_1} \cdot {x_2}^{k_2} \cdots {x_P}^{k_P} ,$$

taken over all natnum-tuples $\vec{k}=(k_1,\ldots,k_P)$ that sum to N.

Preliminaries. Fix a posint N and let "perm" mean a permutation of $[1..N]$. Given an N-variable poly $Y(z_1,\ldots,z_N)$ and a perm π, say that Y is ***π-invariant*** if

$$Y(z_1,\ldots,z_N) = Y(z_{\pi(1)}, z_{\pi(2)}, \ldots, z_{\pi(N)}) ,$$

where the equality is "as polynomials"; that is, *corresponding coefficients are equal*. Say that Y is a "***symmetric polynomial*** in $z_1,\ldots,z_N$" if Y is invariant[♡1] under each of the $N!$ many perms. I will use ***sympoly*** to abbreviate "symmetric polynomial".

1: *Modification.* For notational convenience, henceforth assume that sympolys have no constant term. This does not affect the below ESP Theorem, (3), since the constant of $Y()$ would simply become the constant term of $F()$. □

ESP. In the sequel, use $\vec{z}$ to denote $z_1,\ldots,z_N$ or $(z_1,\ldots,z_N)$ as appropriate.

For N=3, there are three especially simple sympolys:

$$\begin{aligned} \sigma_1(\vec{z}) &:= z_1 + z_2 + z_3 ; \\ \sigma_2(\vec{z}) &:= z_1 z_2 + z_1 z_3 + z_2 z_3 ; \\ \sigma_3(\vec{z}) &:= z_1 \cdot z_2 \cdot z_3 . \end{aligned}$$

For a general N and $j \in [1..N]$, let

2: $$\sigma_{j,N}(\vec{z}) := \sum_S \Big[\prod_{j \in S} z_j\Big] ,$$

where S, here, ranges over the cardinality-j subsets of $[1..N]$. Henceforth, the subscript "N" will be implicit and I will write σ_j in place of $\sigma_{j,N}$.

These sympolys $\sigma_1,\ldots,\sigma_N$ are called the ***elementary symmetric polynomials***, or ***esp***s, for short.

The goal of this note is the following result.

[♡1] Consider the polynomial $Y(x,y) := x^2 + y$ as a poly over the integers mod-2. Then Y is certainly a symmetric *function*, but it is not a "symmetric polynomial" because its coefficient structure is not invariant under exchanging x and y.

3: Fundamental ESP Theorem. *Fix an N-variable $\mathbb{Z}$-sympoly Y. Then there is a unique N-variable $\mathbb{Z}$-poly F such that*

$$3': \quad F\Big(\sigma_1(\vec{z}), \sigma_2(\vec{z}), \ldots, \sigma_N(\vec{z})\Big) \;=\; Y(\vec{z})\,.$$

Additionally, $\operatorname{Deg}(F) \leqslant \operatorname{Deg}(Y)$. ◊

Convention. Typically I will let σ_j also mean $\sigma_j(\vec{z})$. In particular, I will usually write $\text{LhS}(3')$ as $F(\sigma_1, \ldots, \sigma_N)$ or just as $F(\vec{\sigma})$.

When I use a specific value of N in an example, I may use variable names $a, b, \ldots$ in place of $z_1, z_2, \ldots$. □

Remark. With alternating signs, the coefficients of a monic 1-var poly $g()$ appear as the esps of its zeros $\vec{z}$ (also called "roots" of g). I.e, consider

$$g(X) \;:=\; [X - z_1][X - z_2]\cdots[X - z_N]\,.$$

Multiplying this out, $g(X)$ equals

$$X^N - \sigma_1(\vec{z})\cdot X^{N-1} + \sigma_2(\vec{z})\cdot X^{N-2} \\ - \sigma_3(\vec{z})\cdot X^{N-3} + \cdots + [\text{-}1]^N \sigma_N(\vec{z})\cdot X^0\,.$$

Consequently, the ESP Theorem has the handy corollary that each symmetric polynomial of the roots of $g()$ is some nice polynomial in the *coefficients* of g.

As an application, the "***discriminant*** of g" is

$$4: \qquad \operatorname{Discr}(g) \;:=\; \prod_{1\leqslant j<k\leqslant N} [z_k - z_j]^2\,.$$

Letting B be the binomial coeff $\binom{N}{2}$, the RhS equals

$$[\text{-}1]^B \cdot \prod_{\substack{j\neq k \\ j,k\in[1\,..\,N]}} [z_k - z_j]\,,$$

so we see that $\operatorname{Discr}(g)$ does *not* depend on some arbitrary ordering of the roots. The ESP Thm tells us that we can compute the discriminant of g from its coefficients. □

Tools

Here are the utensils that we will use to obtain the ESP Theorem.

Dictionary order. We define ***dictionary order***, $\prec$, on pairs of N-tuples $\vec{v} = (v_1, \ldots, v_N)$ and $\vec{w}$, each a tuple of natnums. Say that $\vec{v} \prec \vec{w}$ if:

> *There exists an index $j \in [1\,..\,N]$ with $v_j \neq w_j$. For the smallest such j, furthermore, $v_j < w_j$.*

Easily, dictionary order[♡2] is *transitive* ($\vec{u} \preccurlyeq \vec{v}$ and $\vec{v} \preccurlyeq \vec{w}$ together imply that $\vec{u} \preccurlyeq \vec{w}$). And dictionary order is a *total* order. That is, for each pair $\vec{v}$ and $\vec{w}$, one of them is $\preccurlyeq$ the other one.

Profiles. A sympoly can be written more concisely than just listing all of its terms. Consider the N=3 sympoly $Y()$ with the smallest number of terms and possessing term $a^9 \cdot b^4 \cdot c$. Necessarily $Y(a,b,c)$ equals this sum of six terms:

$$a^9b^4c + a^9bc^4 + a^4b^9c + a^4bc^9 + ab^9c^4 + ab^4c^9$$

What we have done is take the exponent triple $(9,4,1)$ and put it over (a,b,c) in all possible ways.

In general, an exponent N-tuple, $\boldsymbol{\alpha}$, can always be put in decreasing (well…non-increasing, actually) order as $\boldsymbol{\alpha} = (h_1, \ldots, h_N)$ with $h_1 \geqslant \ldots \geqslant h_N$, with each of the exponents a natnum. Furthermore, courtesy (1), necessarily $h_1 \geqslant 1$. We take this as our definition of a ***profile***.

Let's refer to the N natnums h_n as ***heights***. They are indeed shown as such in Figure 9, below. Let $\operatorname{Deg}(\boldsymbol{\alpha})$ denote the sum $h_1 + \ldots + h_N$.

The Canonical Rep. Let $S^{\boldsymbol{\alpha}}(\vec{z})$ be the sympoly having the fewest terms and having term $z_1^{h_1}\cdot z_2^{h_2}\cdots z_N^{h_N}$. Letting π, below, range over all permutations of $[1\,..\,N]$, then,

$$5: \quad S^{\boldsymbol{\alpha}}(\vec{z}) \;:=\; \tfrac{1}{B}\cdot\sum_{\pi} z_1^{h_{\pi(1)}}\cdot z_2^{h_{\pi(2)}}\cdots z_N^{h_{\pi(N)}}\,,$$

[♡2] Also known as ***lexicographic order***.

where we need to define the posint B appropriately so that, after combining like-terms, each term will have a coefficient of 1.

To compute B, we count the number of height-repetitions in our N-tuple $\boldsymbol{\alpha}$. Let $v_1, \ldots, v_P$ be the values of $\boldsymbol{\alpha}$, and write

$$\boldsymbol{\alpha} = \Big(\underbrace{v_1, \ldots, v_1}_{d_1}, \underbrace{v_2, \ldots, v_2}_{d_2}, \ldots, \underbrace{v_P, \ldots, v_P}_{d_P}\Big)$$

with strict inequality $v_1 > \cdots > v_P$. (So v_1 equals h_1 and $v_P = h_N$. There are d_j many copies of the j^{th} value.) It is straightforward to show that B equals the product $[d_1!] \cdot [d_2!] \cdots [d_P!]$. Thus $S^{\boldsymbol{\alpha}}$ has a multinomial-coeff number of terms, namely (**??**).

The upshot is this: Using dictionary order on profiles, an arbitrary sympoly $Y()$ has a ***canonical representation*** (abbr. ***canon-rep***)

6: $$Y = \sum_{k=1}^{K} \mathsf{q}_k S^{\boldsymbol{\alpha}_k}\,,$$ where $\boldsymbol{\alpha}_1 \succ \boldsymbol{\alpha}_2 \succ \ldots \succ \boldsymbol{\alpha}_K$ and each q_k is a non-zero integer.

The natnum K, coefficients q_k, and profile sequence $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_K$ are uniquely determined.

Establishing the ESP Theorem

An N-profile can be viewed as in Figure 9. It determines a decreasing tuple of widths $[\![w_1, \ldots, w_L]\!]$. Notice that w_L will always be the number of non-zero heights. Thus necessarily $L \leqslant N$, with equality IFF $\boldsymbol{\alpha}$ is the profile of an ESP. Whoa! Did I mean iff $h_1 = 1$? I.e $L = 1$?

The key step in proving the theorem is this lemma.

7: Width Lemma. *Fix a profile $\boldsymbol{\alpha}$ and let $w_1 \leqslant \ldots \leqslant w_L$ be its tuple of widths. Then there is a posint J so that the esp-product*

$$\sigma_{w_1} \cdot \sigma_{w_2} \cdots \sigma_{w_L}\,, \quad \textit{(evidently a sympoly)}$$

has canonical representation

8: $$\sigma_{w_1} \cdot \sigma_{w_2} \cdots \sigma_{w_L} = S^{\boldsymbol{\alpha}} + \sum_{j=2}^{J} \mathsf{q}_j S^{\boldsymbol{\beta}_j}\,,$$

for some coefficients q_j and profiles $\boldsymbol{\beta}_J \prec \boldsymbol{\beta}_{J-1} \prec \ldots \prec \boldsymbol{\beta}_2 \prec \boldsymbol{\alpha}$. In other words, the original profile $\boldsymbol{\alpha}$ is the highest profile in the esp-product (8), and it occurs with a coefficient of 1.$\diamondsuit$

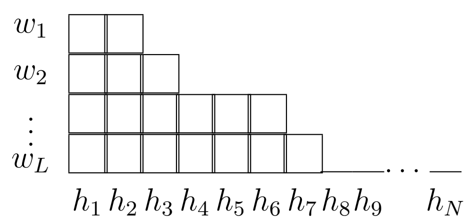

FIG. 9: *This shows a profile $\boldsymbol{\alpha}$ having heights $h_1 \geqslant \ldots \geqslant h_N$. In this example, heights h_8 through h_N are zero. The picture also defines L many widths $w_1 \leqslant \ldots \leqslant w_L$, where $L := h_1$. In this example $L = 4$ and the widths are $w_1 = 2$, $w_2 = 3$, $w_3 = 6$ and $w_4 = 7$.*

10: *Example.* Let $\boldsymbol{\alpha} := (2, 1, 1, 0)$; so $N = 4$. The width tuple for $\boldsymbol{\alpha}$ is $[\![1, 3]\!]$, hence the corresponding esp-product is $\sigma_1 \cdot \sigma_3$, i.e, is

$*$: $$[a + b + c + d][abc + abd + acd + bcd]\,.$$

The *only* Deg =4 profile which is "dictionarily" below $\boldsymbol{\alpha}$ is $\boldsymbol{\beta} := (1, 1, 1, 1)$. Collecting terms in $(*)$ gives that

8′: $$\sigma_1 \cdot \sigma_3 = S^{\boldsymbol{\alpha}} + 4S^{\boldsymbol{\beta}}\,.$$

So (8′) is (8) with $J = 2$ and $\mathsf{q}_2 = 4$.

There is an alternative way to derive (8′), without multiplying out. Courtesy the lemma, we will get one copy of $\boldsymbol{\alpha}$ and some unknown number, q, of copies of $\boldsymbol{\beta}$. But the repetition tuple for $\boldsymbol{\alpha}$ is $(1, 2, 1)$, so $S^{\boldsymbol{\alpha}}$ has $\binom{4}{1,2,1}$=12 terms. Thus 12 of the $16 = 4 \cdot 4$ products of $(*)$ will be terms of the single copy of $S^{\boldsymbol{\alpha}}$.

In consequence, the remaining $16-12 = 4$ products, when divided by the number of terms in $S^{\boldsymbol{\beta}}$, must equal q. Since $S^{\boldsymbol{\beta}}$ has only a single term, *abcd*, we conclude that $\mathsf{q} = 4/1$; this is also what we saw by multiplying out. □

Sketch of proof of the Width Lemma. There is no true loss of generality in assuming that $N = 5$; now I can write the variables as $a, \ldots, e$ rather than $z_1, \ldots, z_5$. Referring to (6), let

$$\text{11:} \qquad \mathsf{p} \cdot S^{\boldsymbol{\alpha}'} + \sum_{k=2}^{K} \mathsf{q}_k S^{\boldsymbol{\beta}_k}$$

be the canon-rep of the esp-product $\sigma_{w_1} \cdot \sigma_{w_2} \cdots \sigma_{w_L}$.

Our goal is to show that $\boldsymbol{\alpha}'$ equals $\boldsymbol{\alpha}$ and that $\mathsf{p} = 1$. It is implicit in the argument below (Exercise!) that no profile dictionarily greater than $\boldsymbol{\alpha}$ can occur in (11). So, I will content myself with showing that the coeff of $\boldsymbol{\alpha}$ in (11) is 1.

The coeff of $\boldsymbol{\alpha}$ is simply the coeff of the term

$$\text{**:} \qquad a^{h_1} b^{h_2} c^{h_3} d^{h_4} e^{h_5}$$

in esp-product $\sigma_{w_1} \cdots \sigma_{w_L}$. In each of the L many esps, the degree of "a" is 1. Thus the highest degree of "a" in the esp-product is L, which indeed equals h_1. How many of the variables $a, b, \ldots$ can have this maximal exponent L? —at most w_1 of them. So the only way to get term (**) in the esp-product is if the term used from esp σ_{w_1} was in fact the term *abcde*.

Rest of proof is missing, 17Sep2001. ♦

An implementation

The algorithm will manipulate lists OTForm

$$\text{12:} \qquad \mathtt{y} = \Big\langle (\boldsymbol{\alpha}_1, \mathsf{c}_1), (\boldsymbol{\alpha}_2, \mathsf{c}_2), \ldots, (\boldsymbol{\alpha}_K, \mathsf{c}_K) \Big\rangle,$$

where $\boldsymbol{\alpha}_1 \succ \ldots \succ \boldsymbol{\alpha}_K$ are profiles and the coeffs c_k are non-zero integers.

Given a second such list

$$\mathtt{x} = \Big\langle (\boldsymbol{\beta}_1, \mathsf{q}_1), \ldots, (\boldsymbol{\beta}_K, \mathsf{q}_K) \Big\rangle,$$

let $\mathtt{merge\text{-}sort}(\mathtt{y}, \mathtt{x})$ be the interwoven list of pairs, sorted by $\succ$. Furthermore, like-terms are combined and terms with coeff zero are dropped. For example, suppose that $\boldsymbol{\alpha} \succ \boldsymbol{\beta} \succ \boldsymbol{\gamma} \succ \boldsymbol{\delta} \succ \boldsymbol{\varepsilon}$ are profiles and our lists are

$$\mathtt{y} := \Big\langle (\boldsymbol{\beta}, 20), (\boldsymbol{\gamma}, \text{-}30), (\boldsymbol{\varepsilon}, 40) \Big\rangle;$$
$$\mathtt{x} := \Big\langle (\boldsymbol{\alpha}, 5), (\boldsymbol{\beta}, 6), (\boldsymbol{\gamma}, 30), (\boldsymbol{\delta}, 7), (\boldsymbol{\varepsilon}, 8) \Big\rangle.$$

Then $\mathtt{merge\text{-}sort}(\mathtt{y}, \mathtt{x})$ equals

$$\Big\langle (\boldsymbol{\alpha}, 5), (\boldsymbol{\beta}, 26), (\boldsymbol{\delta}, 7), (\boldsymbol{\varepsilon}, 48) \Big\rangle.$$

The setup. The given sympoly has canonical representation

$$\text{IN:} \qquad Y = \sum_{k=1}^{K} \mathsf{c}_k S^{\boldsymbol{\alpha}_k}.$$

The desired output poly F has form

$$\text{OUT:} \qquad F = \sum_{j=1}^{J} \mathsf{d}_j \Upsilon_j,$$

where each Υ_j is OTForm $\sigma_{w_1} \cdots \sigma_{w_L}$ for numbers L and $w_1 \leqslant \ldots \leqslant w_L$ which depend on j.

An algorithm

The algorithm repeats *Steps 1,2,3* below until list $\mathtt{y}$ becomes empty. The value of list $\mathtt{f}$ then describes the desired poly F.

Initialization: Init $\mathtt{y}$ to list (12) from (IN). Init $\mathtt{f}$ to the empty list $\langle\rangle$.

Step 1: If the current $\mathtt{y}$ is empty then STOP; the current value of $\mathtt{f}$ describes (OUT).

Otherwise, *pop-off* of $\mathtt{y}$ its leftmost pair and call this pair $(\boldsymbol{\alpha}, \mathsf{c})$. Compute L and widths $w_1 \leqslant \ldots \leqslant w_L$ for profile $\boldsymbol{\alpha}$.

Step 2: Courtesy the Width Lemma, we can write the product $\sigma_{w_1} \cdots \sigma_{w_L}$ as

$$\Big\langle \big(\boldsymbol{\alpha}, 1\big), \big(\boldsymbol{\beta}_2, \mathsf{q}_2\big) \ldots, \big(\boldsymbol{\beta}_J, \mathsf{q}_J\big) \Big\rangle$$

where $\boldsymbol{\alpha} \succ \boldsymbol{\beta}_2 \succ \ldots \succ \boldsymbol{\beta}_J$ are profiles.

Push pair $\Big([w_1, w_2, \ldots, w_L], \mathsf{c}\Big)$ onto `f`.

Step 3: Define a list

$$\mathtt{x} \coloneqq \Big\langle \big(\boldsymbol{\beta}_2, \text{-}\mathsf{c}\mathsf{q}_2\big), \big(\boldsymbol{\beta}_3, \text{-}\mathsf{c}\mathsf{q}_3\big), \ldots, \big(\boldsymbol{\beta}_J, \text{-}\mathsf{c}\mathsf{q}_J\big) \Big\rangle\,;$$

note the negated coefficients. Now assign

$$\mathtt{y} \;\coloneqq\; \mathtt{merge\text{-}sort}(\mathtt{y}, \mathtt{x})\,.$$

(Let K denote the number of pairs in the old `y`. Then the number of pairs in the new `y` may be as much as $K+J-2$. But the maximum profile in the new `y` is strictly $\prec$ the maximum profile in old `y`. So eventually `y` will become exhausted.)

Efficiency considerations

An expensive part is *Step 2*, computing the canon-rep of the esp-product.

To be written.

Filename: `Problems/Polynomials/symmetric_polys.latex`
As of: *Wednesday 15Feb2006.* Typeset: *14Oct2017* at *11:56.*