Math 701 Fall 2026, Lecture 1: How symmetric polynomials appear

website: `https://www.cip.ifi.lmu.de/~grinberg/t/26fs/`

0.1. General

This is a course on symmetric and quasisymmetric functions. We will start with classical results that are well represented in textbooks and lecture notes; then we will gradually move on to more active research topics.

The course website (where all lecture notes and homeworks will be posted) is

`https://www.cip.ifi.lmu.de/~grinberg/t/26fs/`

But homework should be submitted via Canvas. Homework must be done without AI and without consulting outsiders; collaboration is allowed as long as you acknowledge your collaborators. Homework should be typeset, not handwritten.

0.2. Some motivation

Historically speaking, symmetric polynomials/functions first appeared in the study of roots of polynomials, and this has given a major impetus for their early study by Newton and Gauss. Let me show a slight variant on this motivation.

0.2.1. Playing with Fibonacci numbers

Recall the **Fibonacci sequence**: This is the integer sequence $(f_0, f_1, f_2, \ldots)$ defined recursively by $f_0 = 0$ and $f_1 = 1$ and

$$f_n = f_{n-1} + f_{n-2} \qquad \text{for each } n \geq 2. \tag{1}$$

The latter equality $f_n = f_{n-1} + f_{n-2}$ is called a **linear recurrence**, as it expresses each term f_n of the sequence as a linear combination (with constant coefficients) of the previous two terms f_{n-1} and f_{n-2}. (This is a discrete analogue of linear ODEs with constant coefficients, a venerable subject!)

Now consider the sequence $(f_0, f_2, f_4, f_6, \ldots)$, which consists of every second Fibonacci number. I claim that it also satisfies a linear recurrence: namely,

$$f_{2n} = 3f_{2n-2} - f_{2n-4} \qquad \text{for all } n \geq 2.$$

In fact, this is not specific to even-indexed Fibonacci numbers; more generally, we have

$$f_m = 3f_{m-2} - f_{m-4} \qquad \text{for all } m \geq 4, \tag{2}$$

as we can show by applying the original Fibonacci recurrence (1) multiple times:

$$\begin{aligned} f_m &= \underbrace{f_{m-1}}_{\substack{=f_{m-2}+f_{m-3} \\ \text{(by (1))}}} + f_{m-2} \qquad \text{(by (1))} \\ &= (f_{m-2}+f_{m-3}) + f_{m-2} = 2f_{m-2} + \underbrace{f_{m-3}}_{\substack{=f_{m-2}-f_{m-4} \\ \text{(since (1) yields } f_{m-2}=f_{m-3}+f_{m-4})}} \\ &= 2f_{m-2} + f_{m-2} - f_{m-4} = 3f_{m-2} - f_{m-4}. \end{aligned}$$

So not only does the every-second-Fibonacci-number sequence $(f_0, f_2, f_4, f_6, \ldots)$ satisfy a linear recurrence, but its offset variant $(f_1, f_3, f_5, f_7, \ldots)$ satisfies the same linear recurrence.

What about every third Fibonacci number? Does the sequence $(f_0, f_3, f_6, f_9, \ldots)$ satisfy a linear recurrence, too? We can try to find such a recurrence by "following our nose" using (1) again – e.g. as follows:

$$\begin{aligned} f_m &= f_{m-1} + f_{m-2} = 2f_{m-2} + f_{m-3} = 2\left(f_{m-3} + f_{m-4}\right) + f_{m-3} \\ &= 3f_{m-3} + 2f_{m-4} = 3f_{m-3} + 2\left(f_{m-5} + f_{m-6}\right) = \cdots \end{aligned}$$

– but it is not clear how to finish this: we would like to get rid of all appearances of $f_{m-1}, f_{m-2}, f_{m-4}, f_{m-5}, f_{m-7}, \ldots$ on the right-hand side, but it is not obvious how to do so.

Yet it can be shown that the sequence $(f_0, f_3, f_6, f_9, \ldots)$ satisfies a linear recurrence, which is again shared by its two offset variants $(f_1, f_4, f_7, f_{10}, \ldots)$ and $(f_2, f_5, f_8, f_{11}, \ldots)$. Namely, we have

$$f_m = 4f_{m-3} + f_{m-6} \qquad \text{for all } m \geq 6. \tag{3}$$

Proving this is straightforward (e.g., by strong induction), but how was this recurrence found? And can it be generalized? What about every fourth or every fifth Fibonacci number? More generally, what about every m-th Fibonacci number for a given $m \in \mathbb{N}$?

0.2.2. Linear recurrences in general

Actually, what about linearly recurrent sequences of higher order? Let us recall what this means:

Definition 0.2.1. Let k be a positive integer. Let $(x_0, x_1, x_2, \ldots)$ be a sequence of numbers (e.g., complex numbers). We say that this sequence is **linearly recurrent of order** k if it satisfies the recurrence

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \cdots + a_k x_{n-k} \qquad \text{for all } n \geq k, \tag{4}$$

where $a_1, a_2, \ldots, a_k$ are fixed constants (not depending on n). These constants $a_1, a_2, \ldots, a_k$ are called the **coefficients** of this recurrence.

Examples are easy to find:

- The Fibonacci sequence is linearly recurrent of order 2, with coefficients 1 and 1.
- Any geometric progression is linearly recurrent of order 1 (the recurrence being $x_n = qx_{n-1}$).
- Any arithmetic progression is linearly recurrent of order 2 (the recurrence being $x_n = 2x_{n-1} - x_{n-2}$).

Note that a linearly recurrent sequence of order k will also be linearly recurrent of order $k+1$ (just add a "$+0x_{n-k-1}$" term to (4)), so the coefficients belong to the recurrence, not to the sequence.

Now we can generalize the question we posed above for the Fibonacci sequence:

> **Question 0.2.2.** Let $(x_0, x_1, x_2, \ldots)$ be a linearly recurrent sequence of order k. Let $m \in \mathbb{N}$ be arbitrary. Then, is the sequence $(x_0, x_m, x_{2m}, x_{3m}, \ldots)$ again linearly recurrent of order k ? If so, how can we compute the coefficients of the recurrence?

0.2.3. The matrix trick

The main trick involved in linearly recurrent sequences is viewing them as entries of matrix powers. Namely, consider a linearly recurrent sequence $(x_0, x_1, x_2, \ldots)$ of order k with coefficients $a_1, a_2, \ldots, a_k$. Define the $k \times k$-matrix

$$A := \begin{pmatrix} a_1 & a_2 & \cdots & a_{k-1} & a_k \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \tag{5}$$

(where each of the 2-nd, 3-rd, ..., k-th rows has just a single 1 surrounded by zeroes, and these 1's are situated in columns $1, 2, \ldots, k-1$ from top to bottom).

For each $n \geq 0$, define the vector

$$v_n := \begin{pmatrix} x_{n+k-1} \\ x_{n+k-2} \\ \vdots \\ x_n \end{pmatrix};$$

this consists of k consecutive entries of our sequence, ending with x_n. Then, our recurrence equation (4) yields

$$v_n = Av_{n-1} \qquad \text{for each } n \geq 1, \tag{6}$$

since

$$v_n = \begin{pmatrix} x_{n+k-1} \\ x_{n+k-2} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_1 x_{n+k-2} + a_2 x_{n+k-3} + \cdots + a_k x_{n-1} \\ x_{n+k-2} \\ \vdots \\ x_n \end{pmatrix}$$

(here, we rewrote the topmost entry using (4))

$$= \underbrace{\begin{pmatrix} a_1 & a_2 & \cdots & a_{k-1} & a_k \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} x_{n+k-2} \\ x_{n+k-3} \\ \vdots \\ x_{n-1} \end{pmatrix}}_{=v_{n-1}} = Av_{n-1}.$$

Applying this equality repeatedly, we find

$$v_n = A \underbrace{v_{n-1}}_{=Av_{n-2}} = AA \underbrace{v_{n-2}}_{=Av_{n-3}} = AAAv_{n-3} = \cdots = A^n v_0 \qquad \text{for each } n \geq 0.$$

Thus,

$$v_n = A^n v_0 \qquad \text{for each } n \geq 0. \tag{7}$$

And this makes it easy to compute v_n (and therefore x_n, which is the last entry of v_n) by taking the matrix A to the n-th power. This is actually a useful observation in computing high values of linearly recurrent sequences, since powers of matrices can be computed using fairly efficient algorithms (exponentiation by squaring).

But we are interested in the sequence $(x_0, x_m, x_{2m}, x_{3m}, \ldots)$. This sequence consists of the last entries of the vectors $v_0, v_m, v_{2m}, v_{3m}, \ldots$ (since each x_n is the last entry of v_n). We want to know whether it is linearly recurrent of order k. Thus we want to find linear relations between the vectors $v_0, v_m, v_{2m}, v_{3m}, \ldots$. Using (7), we can rewrite these vectors as $A^0 v_0$, $A^m v_0$, $A^{2m} v_0$, $A^{3m} v_0$, ... Therefore, if we have a linear relation between the matrices $A^0, A^m, A^{2m}, A^{3m}, \ldots$, then (by multiplying with the vector v_0) we will obtain a linear relation between the vectors $v_0, v_m, v_{2m}, v_{3m}, \ldots$, and then (by taking the last entries) also between the numbers $x_0, x_m, x_{2m}, x_{3m}, \ldots$.

A well-known source of linear relations between powers of a matrix is the **Cayley–Hamilton theorem**. This theorem (which is proved in any good linear algebra text) says that if B is any $k \times k$-matrix, and if $\chi_B(t) := \det(tI_k - B)$ is its characteristic polynomial[1] (in the indeterminate t), then

$$\chi_B(B) = 0. \tag{8}$$

[1]Some authors define the characteristic polynomial to be $\det(B - tI_k)$ instead. This differs from our definition only by a sign.

Recall that χ_B is always a monic polynomial of degree k, that is, a polynomial of the form $\chi_B(t) = t^k - c_1 t^{k-1} - c_2 t^{k-2} - \cdots - c_k t^0$ for some numbers $c_1, c_2, \ldots, c_k$. Writing χ_B in this form, we can rewrite (8) as

$$B^k - c_1 B^{k-1} - c_2 B^{k-2} - \cdots - c_k B^0 = 0, \tag{9}$$

that is, as

$$B^k = c_1 B^{k-1} + c_2 B^{k-2} + \cdots + c_k B^0. \tag{10}$$

We can apply this to $B = A^m$, and obtain

$$A^{mk} = c_1 A^{m(k-1)} + c_2 A^{m(k-2)} + \cdots + c_k A^0, \tag{11}$$

where $c_1, c_2, \ldots, c_k$ are constant numbers such that $\chi_{A^m}(t) = t^k - c_1 t^{k-1} - c_2 t^{k-2} - \cdots - c_k t^0$ (that is, $c_1, c_2, \ldots, c_k$ are the first k coefficients[2] of $\chi_{A^m}(t)$, multiplied by -1). Hence, for each $n \geq mk$, we have

$$A^n = c_1 A^{n-m} + c_2 A^{n-2m} + \cdots + c_k A^{n-km} \tag{12}$$

(this follows from (11) by multiplying both sides by A^{n-mk}). Multiplying both sides of (12) by v_0, we obtain

$$A^n v_0 = c_1 A^{n-m} v_0 + c_2 A^{n-2m} v_0 + \cdots + c_k A^{n-km} v_0.$$

In view of (7), we can rewrite this as

$$v_n = c_1 v_{n-m} + c_2 v_{n-2m} + \cdots + c_k v_{n-km}.$$

Taking the last entries of all the vectors in this equality, we find

$$x_n = c_1 x_{n-m} + c_2 x_{n-2m} + \cdots + c_k x_{n-km} \tag{13}$$

(since the last entry of each vector v_i is x_i).

Thus, we have proved that each $n \geq km$ satisfies the equality (13). That is, we have proved the following:

Theorem 0.2.3. Let $a_1, a_2, \ldots, a_k$ be k numbers. Let $(x_0, x_1, x_2, \ldots)$ be a linearly recurrent sequence of order k with coefficients $a_1, a_2, \ldots, a_k$. Let $m \in \mathbb{N}$. Define a $k \times k$-matrix A by (5). Let $c_1, c_2, \ldots, c_k$ be the first k coefficients of the characteristic polynomial $\chi_{A^m}(t)$, multiplied by -1. Then, we have

$$x_n = c_1 x_{n-m} + c_2 x_{n-2m} + \cdots + c_k x_{n-km} \qquad \text{for each } n \geq km. \tag{14}$$

In particular, the sequence $(x_0, x_m, x_{2m}, x_{3m}, \ldots)$ is linearly recurrent of order k with coefficients $c_1, c_2, \ldots, c_k$; more generally, for each $i \in \mathbb{N}$, the sequence $(x_i, x_{i+m}, x_{i+2m}, x_{i+3m}, \ldots)$ is linearly recurrent with the same coefficients.

[2] By "the first k coefficients", we mean the coefficients of the monomials $t^0, t^1, \ldots, t^{k-1}$.

For example, let us apply this theorem to the Fibonacci sequence $(f_0, f_1, f_2, \ldots)$ and to $m = 2$. Here, we have $k = 2$ and $a_1 = 1$ and $a_2 = 1$, so that $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$. Thus, $A^2 = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, which has characteristic polynomial

$$\chi_{A^2}(t) = \det\left(tI_2 - \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} t-2 & -1 \\ -1 & t-1 \end{pmatrix} = t^2 - 3t + 1.$$

Thus, when we apply Theorem 0.2.3, we must set $c_1 = 3$ and $c_2 = -1$ (corresponding to the coefficients -3 and 1 of $\chi_{A^2}(t)$). Hence, (14) says that

$$f_n = 3f_{n-2} + (-1) f_{n-4} \qquad \text{for each } n \geq 2 \cdot 2.$$

This is precisely the equality (2) that we found above. Likewise, we can find (3) and similar recurrences for higher values of m. More generally, this shows that the answer to Question 0.2.2 is: Yes, all these sequences are linearly recurrent – even better: linearly recurrent, all with the same coefficients –, and yes, these coefficients can be computed.

0.2.4. Computing χ_{A^m} from χ_A ?

Now let us try to improve on the above. Theorem 0.2.3 gives **some** algorithm for computing the coefficients $c_1, c_2, \ldots, c_k$ in terms of $a_1, a_2, \ldots, a_k$, but this algorithm is rather cumbersome: first, form the matrix A; then take its power A^m; then expand its characteristic polynomial χ_{A^m}. Isn't there an easier way?

A general fact in linear algebra says that the characteristic polynomial χ_{A^m} (for an arbitrary $k \times k$-matrix A, not necessarily the one given by (5)[3]) is completely determined by m and χ_A; we don't have to know the matrix A itself. The easiest way to see this is using eigenvalues: If the $k \times k$-matrix A has eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ (listed with algebraic multiplicity), then A^m has eigenvalues $\lambda_1^m, \lambda_2^m, \ldots, \lambda_k^m$ (by the spectral mapping theorem[4]). The eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ of A are precisely the roots of χ_A, so that we have the polynomial identity

$$\chi_A(t) = (t - \lambda_1)(t - \lambda_2) \cdots (t - \lambda_k), \tag{15}$$

which shows that χ_A both determines and is determined by $\lambda_1, \lambda_2, \ldots, \lambda_k$. Hence,

$$\chi_{A^m}(t) = (t - \lambda_1^m)(t - \lambda_2^m) \cdots (t - \lambda_k^m), \tag{16}$$

since A^m has eigenvalues $\lambda_1^m, \lambda_2^m, \ldots, \lambda_k^m$. We can theoretically use this formula to compute χ_{A^m} from χ_A, since $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the roots of χ_A.

[3]For the matrix A given in (5), the situation is particularly nice, since we can show that

$$\chi_A(t) = t^k - a_1 t^{k-1} - a_2 t^{k-2} - \cdots - a_k t^0$$

(exercise!), which renders the computation of χ_A trivial. But everything we will be doing below applies to any $k \times k$-matrix A.

[4]If you don't know this fact, prove it, e.g., using Schur triangulation.

But this is even less of a practically useful algorithm than the one in Theorem 0.2.3, because the roots of a general polynomial cannot be computed algebraically! (At least not in the naive "quadratic formula" sense; see the Abel–Ruffini theorem.) Numerical algorithms exist, but we don't want to lose precision.

So what we ideally want is a way to find the coefficients of χ_{A^m} in terms of the coefficients of χ_A without having to factor χ_A. To this aim, we first expand (15):

$$\begin{aligned}
\chi_A(t) &= (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_k) \\
&= t^k - (\lambda_1+\lambda_2+\cdots+\lambda_k)\,t^{k-1} \\
&\qquad + \underbrace{(\lambda_1\lambda_2+\lambda_1\lambda_3+\cdots+\lambda_{k-1}\lambda_k)}_{=\sum_{i<j}\lambda_i\lambda_j}\,t^{k-2} \\
&\qquad - \underbrace{(\lambda_1\lambda_2\lambda_3+\lambda_1\lambda_2\lambda_4+\cdots+\lambda_{k-2}\lambda_{k-1}\lambda_k)}_{=\sum_{i<j<\ell}\lambda_i\lambda_j\lambda_\ell}\,t^{k-3} \\
&\qquad \pm\cdots \\
&\qquad + (-1)^k\,\lambda_1\lambda_2\cdots\lambda_k \\
&= \sum_{i=0}^{k}(-1)^i\,e_i(\lambda_1,\lambda_2,\ldots,\lambda_k)\,t^{k-i},
\end{aligned} \tag{17}$$

where we set

$$\begin{aligned}
e_i(\lambda_1,\lambda_2,\ldots,\lambda_k) &:= \sum_{1\le u_1<u_2<\cdots<u_i\le k}\lambda_{u_1}\lambda_{u_2}\cdots\lambda_{u_i} \\
&= \sum_{\substack{L\subseteq[k];\\ |L|=i}}\prod_{\ell\in L}\lambda_\ell \qquad (\text{where } [k]:=\{1,2,\ldots,k\})
\end{aligned}$$

(in particular, $e_k(\lambda_1,\lambda_2,\ldots,\lambda_k)=\lambda_1\lambda_2\cdots\lambda_k$ and $e_0(\lambda_1,\lambda_2,\ldots,\lambda_k)=1$, because the empty product is 1). This latter number $e_i(\lambda_1,\lambda_2,\ldots,\lambda_k)$ is called the ***i*-th elementary symmetric polynomial** in the $\lambda_1,\lambda_2,\ldots,\lambda_k$. So the formula (17) says that the coefficients of χ_A are precisely these $e_i(\lambda_1,\lambda_2,\ldots,\lambda_k)$, up to sign.

Therefore, in order to compute χ_{A^m} from χ_A, we just need to compute the $e_i(\lambda_1^m,\lambda_2^m,\ldots,\lambda_k^m)$ from the $e_i(\lambda_1,\lambda_2,\ldots,\lambda_k)$. This is best done algebraically, without solving for the $\lambda_1,\lambda_2,\ldots,\lambda_k$. To do so, we forget about what the λ_is are, and just study the elementary symmetric polynomials of any k inputs.

Before we do this systematically, let us do the cases $(m,i)=(2,1)$ and $(m,i)=(2,2)$ by hand. So we want formulas for $e_1(\lambda_1^2,\lambda_2^2,\ldots,\lambda_k^2)$ and $e_2(\lambda_1^2,\lambda_2^2,\ldots,\lambda_k^2)$.

The former is easy:

$$\begin{aligned}
& e_1\left(\lambda_1^2, \lambda_2^2, \ldots, \lambda_k^2\right) \\
& = \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_k^2 \\
& = \left(\underbrace{\lambda_1 + \lambda_2 + \cdots + \lambda_k}_{=e_1(\lambda_1, \lambda_2, \ldots, \lambda_k)}\right)^2 - 2 \underbrace{\sum_{i<j} \lambda_i \lambda_j}_{=e_2(\lambda_1, \lambda_2, \ldots, \lambda_k)} \\
& \qquad \left(\begin{array}{c}\text{by the well-known formula} \\ (\lambda_1 + \lambda_2 + \cdots + \lambda_k)^2 = \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_k^2 + 2\sum_{i<j} \lambda_i \lambda_j\end{array}\right) \\
& = \left(e_1\left(\lambda_1, \lambda_2, \ldots, \lambda_k\right)\right)^2 - 2e_2\left(\lambda_1, \lambda_2, \ldots, \lambda_k\right).
\end{aligned}$$

Computing $e_2\left(\lambda_1^2, \lambda_2^2, \ldots, \lambda_k^2\right)$ is harder with this method, however.