Math 235 Fall 2024, Lecture 6 stenogram: Sums and sequences

website: `https://www.cip.ifi.lmu.de/~grinberg/t/24f`

1. Sums and sequences (cont'd)

1.1. Linear recurrences (cont'd)

Last time, we were trying to find an explicit formula for all entries x_n of an (a,b)-recurrent sequence $(x_0, x_1, x_2, \ldots)$ of numbers. Recall that "(a,b)-recurrent" means that $x_n = ax_{n-1} + bx_{n-2}$ for all $n \geq 2$. We found the formula

$$x_n = \frac{x_1 - \mu x_0}{\lambda - \mu} \lambda^n + \frac{\lambda x_0 - x_1}{\lambda - \mu} \mu^n,$$

where

$$\lambda = \frac{a + \sqrt{a^2 + 4b}}{2}, \qquad \mu = \frac{a - \sqrt{a^2 + 4b}}{2}.$$

But this worked only if $a^2 + 4b \neq 0$. What about the other case?

So let us assume that $a^2 + 4b = 0$. In this case, $b = \dfrac{-a^2}{4}$. Since our above method does not work here, we just look at the first entries of our sequence and see if we spot any patterns:

$$\begin{aligned}
x_0 &= x_0, \\
x_1 &= x_1 = a^0 \left(x_1 - 0x_0\right), \\
x_2 &= ax_1 + bx_0 = ax_1 + \frac{-a^2}{4} x_0 = a\left(x_1 - \frac{a}{4} x_0\right), \\
x_3 &= ax_2 + bx_1 = a\left(ax_1 + \frac{-a^2}{4} x_0\right) + \frac{-a^2}{4} x_1 \\
&= \frac{3a^2}{4} x_1 + \frac{-a^3}{4} x_0 = a^2 \left(\frac{3}{4} x_1 - \frac{a}{4} x_0\right), \\
x_4 &= ax_3 + bx_2 = a\left(\frac{3a^2}{4} x_1 + \frac{-a^3}{4} x_0\right) + \frac{-a^2}{4}\left(ax_1 + \frac{-a^2}{4} x_0\right) \\
&= \frac{a^3}{2} x_1 + \frac{-3a^4}{16} x_0 = a^3 \left(\frac{1}{2} x_1 - \frac{3a}{16} x_0\right), \\
x_5 &= a^4 \left(\frac{5}{16} x_1 - \frac{a}{8} x_0\right) \qquad \text{(similarly)}, \\
x_6 &= a^5 \left(\frac{3}{16} x_1 - \frac{5a}{64} x_0\right), \qquad \text{etc.}
\end{aligned}$$

The formulas seem to follow the pattern

$$x_n = a^{n-1}\left(u_n x_1 - v_n a x_0\right) \qquad \text{for some numbers } u_n, v_n,$$

for all $n \geq 1$. So let us make the ansatz

$$x_n = a^{n-1}\left(u_n x_1 - v_n a x_0\right) \qquad \text{for some numbers } u_n, v_n,$$

and try to compute these numbers u_n, v_n. There are two ways to do this:

- Make a table and try to guess the rule.
- Try to see what the recurrence relation $x_n = ax_{n-1} + bx_{n-2}$ says about u_n and v_n after we substitute the formula $x_n = a^{n-1}\left(u_n x_1 - v_n a x_0\right)$ on both sides.

Let us try the former way:

$$\begin{aligned}
(u_1, u_2, u_3, u_4, u_5, u_6) &= \left(1, 1, \frac{3}{4}, \frac{1}{2}, \frac{5}{16}, \frac{3}{16}\right) \\
&= \left(\frac{1}{1}, \frac{2}{2}, \frac{3}{4}, \frac{4}{8}, \frac{5}{16}, \frac{6}{32}\right); \\
(v_1, v_2, v_3, v_4, v_5, v_6) &= \left(0, \frac{1}{4}, \frac{1}{4}, \frac{3}{16}, \frac{1}{8}, \frac{5}{64}\right) \\
&= \left(\frac{0}{2}, \frac{1}{4}, \frac{2}{8}, \frac{3}{16}, \frac{4}{32}, \frac{5}{64}\right).
\end{aligned}$$

So we guess

$$u_n = \frac{n}{2^{n-1}} = \frac{2n}{2^n} \qquad \text{and} \qquad v_n = \frac{n-1}{2^n}.$$

Thus, our ansatz becomes

$$\begin{aligned}
x_n &= a^{n-1}\left(\frac{2n}{2^n} x_1 - \frac{n-1}{2^n} a x_0\right) \\
&= \frac{a^{n-1}}{2^n}\left(2n x_1 - (n-1)\, a x_0\right) \qquad \text{for all } n \geq 1.
\end{aligned}$$

This can be proved straightforwardly by strong induction on n.

Thus, altogether, we have proved:

Theorem 1.1.1 (generalized Binet formula)**.** Let a and b be two numbers. Let $(x_0, x_1, x_2, \ldots)$ be an (a, b)-recurrent sequence. Then:

(a) If $a^2 + 4b \neq 0$, then every $n \in \mathbb{N}$ satisfies

$$x_n = \gamma\lambda^n + \delta\mu^n,$$

where we set

$$\lambda = \frac{a + \sqrt{a^2 + 4b}}{2} \qquad \text{and} \qquad \mu = \frac{a - \sqrt{a^2 + 4b}}{2}$$

and

$$\gamma = \frac{x_1 - \mu x_0}{\lambda - \mu} \qquad \text{and} \qquad \delta = \frac{\lambda x_0 - x_1}{\lambda - \mu}.$$

(b) If $a^2 + 4b = 0$, then every $n \in \mathbb{N}$ satisfies

$$\begin{aligned} x_n &= \frac{a^{n-1}}{2^n} \left(2nx_1 - (n-1)\, ax_0\right) \\ &= \frac{1}{2^n} \left(2na^{n-1}x_1 - (n-1)\, a^n x_0\right) \end{aligned}$$

(where we understand na^{n-1} as 0 for $n = 0$).

This generalizes

- Binet's formula for the Fibonacci numbers:

$$f_n = \frac{1}{\sqrt{5}}\varphi^n - \frac{1}{\sqrt{5}}\psi^n \qquad \text{with } \varphi, \psi = \frac{1 \pm \sqrt{5}}{2};$$

- a similar formula for the Lucas numbers:

$$\ell_n = \varphi^n + \psi^n \qquad \text{with } \varphi, \psi = \frac{1 \pm \sqrt{5}}{2};$$

- other recurrences like this: For example, for a $(1, -1)$-recurrent sequence, you get

$$\lambda = \frac{1 + \sqrt{-3}}{2} \qquad \text{and} \qquad \mu = \frac{1 - \sqrt{-3}}{2}.$$

  These complex numbers λ and μ are 6-th root of unity, i.e., they satisfy $\lambda^6 = 1$ and $\mu^6 = 1$. Thus, the explicit formula $x_n = \gamma\lambda^n + \delta\mu^n$ shows that the sequence $(x_0, x_1, x_2, \ldots)$ is 6-periodic. This is something we have previously shown algebraically.

1.1.1. Two-term recurrences: various properties

Many facts about the Fibonacci sequence can be generalized to arbitrary (a, b)-recurrent sequences. For instance:

- **Generalized Cassini identity (Exercise 4.9.4):** Let a and b be two numbers. Let $(x_0, x_1, x_2, \ldots)$ be an (a, b)-recurrent sequence. Then,

$$x_{n+1}x_{n-1} - x_n^2 = (-b)^{n-1} \left(x_2 x_0 - x_1^2 \right) \qquad \text{for all } n > 0.$$

- **Generalized addition formula (Exercise 4.9.3):** Let a and b be two numbers. Let $(x_0, x_1, x_2, \ldots)$ and $(y_0, y_1, y_2, \ldots)$ be two (a, b)-recurrent sequences such that $x_0 = 0$ and $x_1 = 1$. Then,

$$y_{n+m+1} = b x_n y_m + x_{n+1} y_{m+1}.$$

  Note that this is a double generalization! We have replaced the Fibonacci sequence not by one but by two (a, b)-recurrent sequences. There is even a further generalization, which gets rid of the $x_0 = 0$ and $x_1 = 1$ assumptions but has a more complicated LHS (Exercise 4.9.4 in the notes).

- **Divisibility property (Exercise 4.9.7):** Let a and b be two integers. Let $(x_0, x_1, x_2, \ldots)$ be an (a, b)-recurrent sequence with $x_0 = 0$ and $x_1 = 1$. Then, all $u, v \in \mathbb{N}$ satisfying $u \mid v$ satisfy $x_u \mid x_v$.

- **Generalized binomial coefficient formula (Proposition 4.9.18):** Let a and b be two numbers such that $a \neq 0$. Let $(x_0, x_1, x_2, \ldots)$ be an (a, b)-recurrent sequence with $x_0 = 0$ and $x_1 = 1$. Then,

$$x_{n+1} = \sum_{k=0}^{n} \binom{n-k}{k} a^{n-2k} b^k \qquad \text{for any } n \geq -1.$$

  This is proved by induction in the notes. Note that the $a \neq 0$ condition is only there to prevent the a^{n-2k} term from being undefined when $n - 2k$ is negative; you can just as well throw these addends away.

A neat application of the latter formula is the identity

$$n + 1 = \sum_{k=0}^{n} (-1)^k \binom{n-k}{k} 2^{n-2k}.$$

Indeed, the sequence $(0, 1, 2, \ldots)$ is $(2, -1)$-recurrent (like any arithmetic progression), so that applying the above formula to $x_i = i$ and $a = 2$ and $b = -1$ yields exactly

$$n + 1 = \sum_{k=0}^{n} \binom{n-k}{k} 2^{n-2k} (-1)^k = \sum_{k=0}^{n} (-1)^k \binom{n-k}{k} 2^{n-2k}.$$

1.1.2. Two-term recurrences: the matrix approach

So far, we have been approaching (a,b)-recurrent sequences $(x_0, x_1, x_2, \ldots)$ entry by entry. However, this is not the best way to approach them, since x_i is not determined by the preceding entry x_{i-1} alone. Meanwhile, **two consecutive entries** x_i and x_{i+1} are determined by the preceding two consecutive entries x_{i-1} and x_i. Thus, we should perhaps look at pairs of consecutive entries at a time. Even better, we can encode such pairs as vectors $\begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix}$. Each such vector $\begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix}$ can be computed from the previous such vector $\begin{pmatrix} x_{i-1} \\ x_i \end{pmatrix}$ by

$$\begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix} = \begin{pmatrix} x_i \\ ax_i + bx_{i-1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix} \begin{pmatrix} x_{i-1} \\ x_i \end{pmatrix}.$$

So we have shown:

Proposition 1.1.2. Let a and b be two numbers. Let A be the 2×2-matrix $\begin{pmatrix} 0 & 1 \\ b & a \end{pmatrix}$.

Let $(x_0, x_1, x_2, \ldots)$ be an (a,b)-recurrent sequence of numbers. For each $i \in \mathbb{N}$, define a column vector v_i by $v_i := \begin{pmatrix} x_i \\ x_{i+1} \end{pmatrix}$. Then:

(a) We have $Av_{i-1} = v_i$ for each $i > 0$.

(b) We have $A^n v_i = v_{i+n}$ for each $i, n \in \mathbb{N}$.

(Proof of **(b)** by induction on n is straightforward.)

This proposition allows you to quickly compute v_n (and thus x_n) if you can quickly compute A^n (for example, exponentiation by squaring lets you do that). But it also helps prove theoretic results. For example, the addition formula can be derived from $A^{n+m} = A^n A^m$. For another example, the generalized Binet formula can be proved (and even discovered) by diagonalizing/jordanizing A. Indeed, if we have a diagonalization/jordanization

$$A = TDT^{-1}$$

of a matrix A, then any power of A can be computed by the formula

$$A^n = TD^nT^{-1},$$

and usually D^n is easy to compute (e.g., if D is diagonal, then D^n is obtained from D simply by taking the n-th powers of all diagonal entries). See the notes for details.

1.1.3. Two-term recurrences: odds and ends

A few more remarks about (a, b)-recurrent sequences before we move on:

- The (entrywise) sum of two (a, b)-recurrent sequences is again (a, b)-recurrent. But not the (entrywise) product.
- (a, b)-recurrent sequences exist not only for numbers but for any objects that can be scaled and added. In particular, there are (a, b)-recurrent sequences of polynomials. The most important example of such sequences are the **Chebyshev polynomials of the first kind** $T_0(x)$, $T_1(x)$, $T_2(x)$, ... defined recursively by

  $$T_0(x) = 1, \qquad T_1(x) = x, \qquad \text{and}$$
  $$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x) \qquad \text{for all } n \geq 2.$$

  So they are a $(2x, -1)$-recurrent sequence of polynomials. They have many useful properties; the most well-known one is the fact that

  $$\cos(n\alpha) = T_n(\cos\alpha) \qquad \text{for any angle } \alpha.$$

There is also such a thing as (a, b, c)-recurrent sequences and, more generally, $(a_1, a_2, \ldots, a_k)$-recurrent sequences. For instance, an (a, b, c)-recurrent sequence is a sequence $(x_0, x_1, x_2, \ldots)$ that satisfies

$$x_n = ax_{n-1} + bx_{n-2} + cx_{n-3} \qquad \text{for all } n \geq 3.$$

These sequences are a lot less well-behaved than (a, b)-recurrent sequences; in particular, they have no addition formula, no binomial coefficient formula(?), no divisibility in general, and the generalized Binet formula for them involves roots of higher-degree polynomials, which are famously hard to explicitly describe.

Yet they sometimes appear in nature (even though not as frequently as (a, b)-recurrent ones). For instance, the sequence $(0^2, 1^2, 2^2, 3^2, \ldots)$ is $(3, -3, 1)$-recurrent, i.e., we have

$$n^2 = 3(n-1)^2 - 3(n-2)^2 + 1(n-3)^2 \qquad \text{for all } n.$$

See the notes for a few more properties.

2. The Extremal Principle

In this chapter, we will learn to use one of the simplest tricks in mathematics: When you have a bunch of objects, look at the smallest or the largest among them. This is known as the **extremal principle**. Figuring out the precise meaning of "smallest" and "largest" might be a choice: For example, if you have a bunch of finite sets of integers, like $\{1,3,6\}$ and $\{2,7\}$ and $\{14,15,18\}$, which one is the largest? The one of largest size? (note that there can be several.) The one with largest minimum? The one with largest maximum? The one with largest sum? You will have to make such decisions, and often you will need to pick the right choice to get something useful.

2.1. Existence theorems

Before we see this principle being used, let me recall some theorems that guarantee the existence of the desired extremal objects:

Theorem 2.1.1. Let S be a nonempty finite set of real numbers. Then, S has a minimum (= a smallest element) and a maximum (= a largest element).

Theorem 2.1.2. Let S be a nonempty set of nonnegative integers. Then, S has a minimum (but usually not a maximum).

Theorem 2.1.3. Let S be a nonempty set of integers. Then:

(a) If S has a lower bound (i.e., there is some integer x such that $x \leq s$ for all $s \in S$), then S has a minimum.

(b) If S has an upper bound, then S has a maximum.

Theorem 2.1.4. Let S be a nonempty set of reals. Then:

(a) If S has a lower bound, then S has an infimum (i.e., a greatest lower bound). This infimum is a minimum of S if and only if it belongs to S.

(b) If S has an upper bound, then S has a supremum (i.e., a least upper bound). This supremum is a maximum of S if and only if it belongs to S.

Theorem 2.1.5. Let S be a nonempty set of reals that is closed with respect to the topology on $\mathbb{R}$. Then:

(a) If S has a lower bound, then S has a minimum.

(b) If S has an upper bound, then S has a maximum.

2.2. Applications

We will now see various uses of the Extremal Principle. We begin with a well-known result, restated in a somewhat unusual form:

Theorem 2.2.1. Let $n \in \mathbb{N}$. Then, there is a unique finite subset T of $\mathbb{N}$ such that $n = \sum_{t \in T} 2^t$.

Proof. This is, of course, just the base-2 representation of n, but let us prove it in a different way.

Existence: We proceed by strong induction on n. In the *induction step*, we want to write n as a sum of distinct powers of 2, and we assume that all nonnegative integers smaller than n can be written in this form.

Pick the **largest** power of 2 smaller or equal to n, that is, the **largest** m such that $2^m \leq n$. (This exists because $2^0 < 2^1 < 2^2 < \cdots$.) Then, $n - 2^m$ is a nonnegative integer smaller than n. Thus, by the induction hypothesis, $n - 2^m$ can be written as a sum of distinct powers of 2, that is,

$$n - 2^m = \sum_{t \in T} 2^t \qquad \text{for some finite set } T \subseteq \mathbb{N}.$$

Hence,

$$n = 2^m + \sum_{t \in T} 2^t.$$

We want to show that this is a sum of distinct powers of 2. The only thing we need to check is that $m \notin T$. Indeed, if we had $m \in T$, then we would have

$$n - 2^m = \sum_{t \in T} 2^t \geq 2^m,$$

so that $n \geq 2^m + 2^m = 2 \cdot 2^m = 2^{m+1}$. That is, we would have $2^{m+1} \leq n$. But this would contradict the maximality of m. This contradiction shows that we do have $m \notin T$, and thus we have written n as a sum of distinct powers of 2.

Uniqueness: We must show that no integer can be written as $\sum_{t \in T} 2^t$ for two different finite subsets T of $\mathbb{N}$. In other words, we must prove that

$$\sum_{t \in T} 2^t \neq \sum_{t \in S} 2^t \qquad \text{for any } T \neq S.$$

To prove this, assume the contrary, and pick the smallest possible n that has two distinct such representations. If the largest elements of T and S are equal, then we can just subtract $2^{\max T = \max S}$ from n and get an even smaller counterexample. If they are distinct, then WLOG $\max T > \max S$, and you can leverage this to show that $\sum_{t \in T} 2^t > \sum_{t \in S} 2^t$. See the notes for details. $\square$

> **Exercise 1.** Let n be a positive integer. A lecture is attended by n students. Each student enters the classroom once and leaves it once. Assume that among any three (distinct) students, there are at least two that overlap (i.e., are together in the room at some moment).
>
> The lecturer wants to make an announcement that every student will hear. Prove that the lecturer can pick two moments at which to make this announcement so that each student hears it at least once. (We assume that making the announcement is instantaneous.)

Translating this into mathematical language, we rewrite this as follows:

> **Exercise 2.** Let n be a positive integer. Let $I_1, I_2, \ldots, I_n$ be n nonempty finite closed intervals on the real axis. Assume that for any three distinct numbers $i, j, k \in \{1, 2, \ldots, n\}$, at least two of the intervals I_i, I_j, I_k intersect (i.e., have a nonempty intersection). Prove that there exist two reals a and b such that each of the intervals $I_1, I_2, \ldots, I_n$ contains at least one of a and b.

Solution. Write each interval I_m as $I_m = [a_m, b_m]$ for some two reals a_m and b_m. Let

$$a := \max\{a_m \mid m \in \{1, 2, \ldots, n\}\} \qquad \text{and}$$
$$b := \min\{b_m \mid m \in \{1, 2, \ldots, n\}\}.$$

(In words: a is the time at which the last student to enter enters, while b is the time at which the first student to leave leaves.)

I claim that a and b are as desired – i.e., that each of the intervals $I_1, I_2, \ldots, I_n$ contains at least one of a and b.

To prove this, we assume the contrary. Thus, some interval $I_p = [a_p, b_p]$ contains neither a nor b. Note that $a = a_u$ for some $u \in \{1, 2, \ldots, n\}$, and $b = b_v$ for some $v \in \{1, 2, \ldots, n\}$. Consider these p, u and v.

By assumption, $[a_p, b_p]$ contains neither a nor b. So it does not contain $a = a_u$. Since

$$a_u = a = \max\{a_m \mid m \in \{1, 2, \ldots, n\}\} \geq a_p,$$

this means that $a_u > b_p$ (since otherwise, we would have $a_u \geq a_p$ and $a_u \leq b_p$ and therefore $a_u \in [a_p, b_p]$). Therefore, the interval $[a_p, b_p]$ lies completely to the left of the interval $[a_u, b_u]$ on the real number line.

A similar argument shows that the interval $[a_p, b_p]$ lies completely to the right of the interval $[a_v, b_v]$ on the real number line.

So the three intervals are arranged as follows:

$$[a_v, b_v] \qquad [a_p, b_p] \qquad [a_u, b_u].$$

In particular, no two of these three intervals intersect. But this contradicts our assumption that among any three of our intervals, at least two intersect. Qed.

Remark 2.2.2. The exercise (in its original formulation, involving lecturer and students) can be understood somewhat more strongly: Can the lecturer pick the two moments "on the fly", without knowing when further students will appear or disappear in the future? The above solution works for this stronger version if the number n is known to the lecturer in advance. If not, it does not (how can the professor find a before it is too late?), and something subtler is needed. See the notes for this.

Exercise 3. Let $n \in \mathbb{N}$. Let $F_1, F_2, \ldots, F_n$ be n distinct points in the plane. Let $W_1, W_2, \ldots, W_n$ be n distinct points in the plane. Prove that there is a way to connect each F-point with a W-point by a line segment so that these line segments do not intersect. In other words, prove that there is a bijection $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ such that no two of the n line segments

$$F_1 W_{\sigma(1)}, \quad F_2 W_{\sigma(2)}, \quad \ldots, \quad F_n W_{\sigma(n)}$$

intersect.