Introduction

“If you already know what recursion is, just remember the answer. Otherwise, find someone who is standing closer to Douglas Hofstadter than you are; then ask him or her what recursion is.” – Andrew Plotkin

Though I do not expect that you will appreciate the cultured humor of that quote, I do expect that you are quite excited about recursion and recurrence relations. Recurrence! A concept that is heavily used in discrete mathematics, combinatorics, computer science, and elsewhere in mathematics. At its core, recursion is a thing that refers to itself; you may be familiar with recursive acronyms such as ‘GNU’ (which stands for ‘GNU is Not Unix’), recursive sorting algorithms, perhaps even recursive number theory if you are familiar with Gödel’s famous theorems.

Recurrences may be beautiful, elegant, and such, but are no use unless we can actually solve problems with them! The journey from a cleverly-spotted recurrence relation to a closed-form answer to the problem we were trying to solve is called ‘solving’ a recurrence (no, this session is not going to be this trivial throughout). In this session, we will explore two methods of solving recurrence relations: characteristic equations (also known as ansatz, German for ‘approach’) and generating functions. There will be exercises meant for each method, though of course, you can use any method to solve them as long as you know what you are doing.

Big Words

Big words are essential to the academic discussion of a complex topic. Our first big word, ‘recurrence’, is one we have already used but not defined. A recurrence is an infinite sequence of numbers \( \{a_n\} \) such that \( a_n = f(a_{n-1}, a_{n-2}, \ldots, a_{n-k}) \) for some function \( f : \mathbb{R}^k \rightarrow \mathbb{R} \) and some positive integer \( k \). Then we have our second big word: \( k \) is the ‘order’ of a recurrence.

We call a recurrence ‘linear’ if all the \( a_i \) terms in \( f \) are of degree 1 and none have a variable coefficients; otherwise we use the big word ‘nonlinear’. If all the terms in \( f \) involve a member of the sequence \( a \), then the recurrence is ‘homogeneous’, otherwise it is ‘nonhomogeneous’, which is truly a big word.

Here are some examples:

linear \( a_n = 3a_{n-2} + a_{n-3} + n + 4 \)

nonlinear \( a_n = 3a_{n-2}^2 + a_{n-3} + n + 4 \) because it has a squared term; \( a_n = 3a_{n-2} + (n-1)a_{n-3} + n + 4 \) because a coefficient has a variable

homogeneous \( a_n = 3a_{n-2} + a_{n-3} \)

nonhomogeneous \( a_n = 3a_{n-2} + a_{n-3} + n + 4 \) because of the \( n + 4 \) part
order 3 $a_n = 3a_{n-2} + a_{n-3} + n + 4$
order 1 $a_n = 3a_{n-1} + n + 4$

nonlinear, nonhomogeneous, order $n$ $a_n = 2^n + \sum_{i=1}^{n} ia_{n-i}$ (yeah, these are tricky)

Finally, you may be wondering that if a sequence is defined in terms of itself, and those terms are also defined in terms or previous terms, then where does it all end? Or equivalently, where does it all start? All recurrences have ‘initial conditions’, and since an order $k$ recurrence needs $k$ previous terms in its definition, it must have $k$ initial conditions, usually defined as $a_0, a_1, \ldots, a_{k-1}$. For example, the recurrence $a_n = 3a_{n-2} + a_{n-3} + n + 4$ may have initial conditions $a_0 = 3$, $a_1 = -1$, and $a_2 = 2$, and that will be enough to get the infinite sequence going.

Characteristic Equations or Ansatz

(Disclaimer: This is the easier method of the two, but can only be used for linear recurrences of constant, nonvariable order.)

Something Simple

This method is best illustrated by an example. Let us consider the famous Fibonacci recurrence: $a_0 = 1$, $a_1 = 1$, and $a_n = a_{n-1} + a_{n-2}$ for all $n \geq 2$. This is linear, homogeneous, and of order 2.

A key step in this method is that we assume $a_n$ is exponential. Indeed, a graph of the Fibonacci numbers appears to rise exponentially. This explains the term ansatz because we take a guessed ‘approach’ at the recurrence, and find that we are right when we are done. Thus, let us write $a_n \approx \alpha^n$ for some $\alpha \neq 0$. We have

$$a_n = a_{n-1} + a_{n-2}
\Rightarrow \alpha^n = \alpha^{n-1} + \alpha^{n-2}
\Rightarrow 0 = \alpha^2 - \alpha - 1$$

Equation (1) is called the ‘characteristic equation’ of the recurrence. We find that it has more than one valid $\alpha$; we will write $a_n$ as a linear combination of all of them. Note that this is another guessed ansatz step that turns out to be right in the end.

We then have $\alpha_1 = \frac{1 + \sqrt{5}}{2}$ and $\alpha_2 = \frac{1 - \sqrt{5}}{2}$, which give $a_n = c_1 \alpha_1^n + c_2 \alpha_2^n$ for some $c_1$ and $c_2$. For specific values, we use our initial conditions to write

$$a_0 = c_1 \alpha_1^0 + c_2 \alpha_2^0
\Rightarrow 0 = c_1 + c_2$$

$$a_1 = c_1 \alpha_1^1 + c_2 \alpha_2^1
\Rightarrow 1 = c_1 \cdot \frac{1 + \sqrt{5}}{2} + c_2 \cdot \frac{1 - \sqrt{5}}{2}$$

This system of linear equations gives $c_1 = \frac{1}{\sqrt{5}}$ and $c_2 = -\frac{1}{\sqrt{5}}$. Finally we have a solution to the Fibonacci recurrence:

$$a_n = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right)$$
A Little Harder

To solve a nonhomogeneous recurrence, we first convert it to a homogeneous recurrence by removing the nonhomogeneous part and solve that recurrence. Then, we yet again guess a form for the nonhomogeneous part and plug that in to the original nonhomogeneous recurrence to obtain a solution for the nonhomogeneous part. We add the homogeneous and nonhomogeneous solutions together to get our final solution.

Before we can apply this method to nonhomogeneous recurrences, there are two caveats we must consider (whose proofs are beyond the scope of this session): firstly, if in solving the characteristic equation we obtain a repeated root $\alpha$ with multiplicity $m$, then for the linear combination of roots, we must use $c_0\alpha^n$, $c_1n\alpha^n$, $c_2n^2\alpha^n$, $\ldots$, $cm^{-1}n^{m-1}\alpha^n$. Other non-repeated roots are used as usual. Secondly, if the nonhomogeneous part of the recurrence is $P(n)s^n$ for some number $s$ and polynomial $P$, then the corresponding nonhomogeneous part in the solution will be $Q(n)n^{m}s^n$ for a polynomial $Q$ with the same degree as $P$, where $m$ is the multiplicity of $s$ in the characteristic equation; $m$ can be zero if $s$ is not a root of the characteristic equation.

Let’s see both of these in practice. We want to solve

$$a_n = 3a_{n-1} - 4a_{n-3} + n3^n$$

(2)

with initial conditions $a_0 = 1$, $a_1 = 1$, $a_2 = 0$. Like Equation (1), we have the characteristic equation

$$\alpha^3 - 3\alpha^2 + 4 = n3^n$$

For the homogeneous part, we remove the $n3^n$ and we are left with $\alpha^3 - 3\alpha^2 + 4 = 0$, which has a double root at $\alpha = 2$ and a single root at $\alpha = -1$. The homogeneous part of the solution will then be $c_02^n + c_1n2^n + c_2(-1)^n$.

For the nonhomogeneous part, our guessed ansatz is $(pm + q)3^n$ according to the second caveat mentioned above. This should satisfy the original recurrence in Equation (2), so

$$(pn + q)3^n = 3(p(n - 1) + q)3^{n-1} - 4(p(n - 3) + q)3^{n-3} + n3^n$$

$$\Rightarrow 3^3pn + 3^3q = 3^3pn - 3^3p + 3^3q - 4pm + 12p - 4q + 3^3n$$

$$\Rightarrow 0 = -27p - 4pm + 12p - 4q + 27n$$

$$= n(27 - 4p) - (15p + 4q)$$

Recall that this must hold for all $n$. This means that both parenthesized terms must be zero, and we have a system of linear equations which gives $p = \frac{27}{4}$ and $q = -\frac{405}{16}$. The nonhomogeneous part of the solution will be $(\frac{27n}{4} - \frac{405}{16})3^n$.

Adding the homogeneous and nonhomogeneous parts,

$$a_n = c_02^n + c_1n2^n + c_2(-1)^n + \left(\frac{27n}{4} - \frac{405}{16}\right)3^n$$

As with the Fibonacci sequence, we use our initial conditions to obtain a system of linear equations:

$$\begin{align*}
1 &= c_0 + c_2 - \frac{405}{16} \\
1 &= 2c_0 + 2c_1 - c_2 + 3\left(\frac{27}{4} - \frac{405}{16}\right) \\
0 &= 4c_0 + 8c_1 + c_2 + 9\left(\frac{27}{4} - \frac{405}{16}\right)
\end{align*}$$

$$\Rightarrow \begin{cases}
c_0 = 28 \\
c_1 = \frac{1}{2} \\
c_2 = -\frac{27}{16}
\end{cases}$$
We finally have the solution to the recurrence in Equation (2) as

\[ a_n = 28 \cdot 2^n - \frac{n}{2} \cdot 2^n - \frac{27}{16} \cdot (-1)^n + \left( \frac{27n}{4} - \frac{405}{16} \right) 3^n \]

(Yes, I checked with WolframAlpha to make sure.)

Problems

For each of these problems, try to first come up with a recurrence that models the situation, then try to solve that recurrence using whatever technique (ansatz, induction, guess-and-check) you see fit. Remember that you must be able to explain and prove your solution semi-rigorously.

1. (inspired by INMO 2015 #4) The employees of FailingProduct, Inc. are playing “Pass the Blame”. The boss has the blame by default. He blames one of his 6 employees. That employee blames another employee (maybe the boss again). After the blame has been passed 8 times, the boss finds himself with the blame. In how many ways can this game of “Pass the Blame” have been played?

2. (Singapore IMO Practice Set) In how many ways can you tile a $2 \times n$ rectangle by $1 \times 1$ squares and L trominoes? (An L tromino is an L-shaped piece with three squares.)

3. (Hong Kong PSC) Find the number of 10-digit positive integers such that each digit is either 1 or 2, and there are two consecutive 1s.

4. (Mercer County Math Circle) Find the units digit of $(3 + \sqrt{7})^{2014}$.

Generating Functions

Generating functions are a very powerful tool. In the words of Herbert S. Wilf who wrote an entire 230-page book\(^1\) about generating functions:

“Generating functions are a bridge between discrete mathematics, on the one hand, and continuous analysis (particularly complex variable theory) on the other. It is possible to study them solely as tools for solving discrete problems. As such there is much that is powerful and magical in the way generating functions give unified methods for handling such problems. … The full beauty of the subject of generating functions emerges from [both] the discrete and the continuous [aspects]. See how they make the solution of difference equations into child’s play. Then see how the theory of functions of a complex variable gives, virtually by inspection, the approximate size of the solution.”

The book is, in fact, recommended reading if you want to go beyond what we will cover here.

What is a generating function? A generating function is the sum of a power series whose coefficients are terms of the sequence we are interested in. For example, if our sequence $a_n$ was 1, 1, 2, 3, 5, 8, … then the power series would be $1x^0, 1x^1, 2x^2, 3x^3, 5x^4, 8x^5, \ldots$ and the sum would be $\sum_{n=0}^{\infty} a_n x^n$. Note that this is just a formal, abstract, algebraic sum; we do not require it to converge or even to actually have some $x$ plugged in. This means we can add, multiply, index-shift, differentiate (!), or do a lot of other things with generating functions without considering their ‘values’ at all.

\(^1\)Generatingfunctionology, available at https://www.math.upenn.edu/~wilf/DownldGF.html.
We introduce two lemmas regarding simple generating functions. The first is merely the geometric sequence sum formula

\[
\frac{1}{1-rx} = 1 + rx + r^2x^2 + r^3x^3 + \cdots
\]  

(3)

Note that while Equation (3) holds only for \(|x| < 1\), we do not really care about \(x\) in a generating function. The second lemma is

\[
\frac{1}{(1-x)^m} = 1 + \binom{m}{1}x + \binom{m+1}{2}x^2 + \binom{m+2}{3}x^3 + \cdots
\]  

(4)

To see why Equation (4) holds, consider the definition of \(\binom{n}{k}\) as \(\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\). Nothing really prevents us from using a negative value of \(n\), so we can write

\[
\binom{-n}{k} = \frac{(-n)(-n-1)(-n-2)\cdots(-n-k+1)}{k!}
\]

\[
= (-1)^k \cdot \frac{n(n+1)(n+2)\cdots(n+k-1)}{k!}
\]

\[
= (-1)^k \binom{n+k-1}{k}
\]

In a similar fashion, we abuse the binomial formula to extend it to negative exponents; as the upper limit of the sum would be lower than the lower limit, we merely sum to infinity (a rigorous proof of why this works is beyond the scope of this session). Then

\[
\frac{1}{(1-x)^m} = (1-x)^{-m} = \sum_{n=0}^{\infty} \binom{-m}{n}(-x)^n
\]

\[
= \sum_{n=0}^{\infty} \binom{m+n-1}{n}x^n
\]

which is the result in Equation (4).

With these lemmas in hand, let us consider Equation (2) and solve it with generating functions. Our generating function will be \(G(x) = \sum_{n=0}^{\infty} a_n x^n\).

\[
a_n = 3a_{n-1} - 4a_{n-3} + n3^n
\]

\[
\Rightarrow \sum_{n=3}^{\infty} a_n x^n = \sum_{n=3}^{\infty} 3a_{n-1} x^n - 4\sum_{n=3}^{\infty} a_{n-3} x^n + \sum_{n=3}^{\infty} n3^n x^n
\]

\[
\Rightarrow G(x) - a_0 - a_1 x - a_2 x^2 = 3x \sum_{n=2}^{\infty} a_n x^n - 4x^3 \sum_{n=0}^{\infty} a_n x^n + \sum_{n=3}^{\infty} n3^n x^n
\]

\[
= 3x(G(x) - a_0 - a_1 x) - 4x^3 G(x) + \sum_{n=3}^{\infty} n3^n x^n
\]

\[
\Rightarrow G(x)(1 - 3x + 4x^3) = (a_2 - 3a_1)x^2 + (a_1 - 3a_0)x + a_0 + \sum_{n=3}^{\infty} n3^n x^n
\]
and using our initial conditions

\[-3x^2 - 2x + 1 + \sum_{n=3}^{\infty} 3^n x^n\]

Combining the results in Equation (3) with \(r = 3\) and Equation (4) with \(m = 2\), and performing a little algebra to get our indexes right, we can write \(\sum_{n=3}^{\infty} n3^nx^n = \frac{3x}{(1-3x)^2} - 3x - 18x^2\). We now write

\[G(x)(1 - 3x + 4x^3) = -21x^2 - 5x + 1 + \frac{3x}{(1 - 3x)^2}\]

\[\implies G(x) = \frac{1 - 5x - 21x^2}{1 - 3x + 4x^3} + \frac{3x}{(1 - 3x)^2(1 - 3x + 4x^3)}\]

\[= \frac{1 - 5x - 21x^2}{(2x - 1)^2(x + 1)} + \frac{3x}{(1 - 3x)^2(2x - 1)^2(x + 1)}\]

Already we can see the roots of the characteristic equation of Equation (2) appear in the denominators as components of the generating function. We can use partial fractions to get (courtesy WolframAlpha)

\[G(x) = \frac{57}{2(1 - 2x)} - \frac{513}{16(1 - 3x)} - \frac{1}{2(1 - 2x)^2} + \frac{27}{4(1 - 3x)^2} + \frac{27}{16(1 + x)}\]

Our two lemmas can convert this partial fraction expansion back into coefficients (remember to get the exponent indices correct!) to give

\[a_n = \left(\frac{57}{2} - \frac{1}{2}\right)2^n - \frac{1}{2} \cdot 2^n - \frac{27}{16} \cdot (-1)^n + \left(\frac{27}{4} - \frac{513}{16} + \frac{27}{4}\right)3^n\]

\[= 28 \cdot 2^n - \frac{n}{2} \cdot 2^n - \frac{27}{16} \cdot (-1)^n + \left(\frac{27}{4} - \frac{405}{16}\right)3^n\]

which is the same result we obtained from the previous method.

**Problems**

For each of these problems, remember that you must be able to explain and prove your solution semi-rigorously. These problems are much more difficult than the previous section. Do not feel ashamed to ask for hints multiple times.

1. Find generating functions for each of

\[\sum_{n=0}^{\infty} nx^{n-1}, \sum_{n=0}^{\infty} nx^n, \sum_{n=0}^{\infty} n^2x^n, \sum_{n=0}^{\infty} (an^2 + bn + c)x^n\]

Find a generating function for \(\sum_{n=0}^{\infty} P(n)x^n\) where \(P\) is a degree-\(m\) polynomial in \(n\).

2. What is the probability of rolling 4 6-sided dice and having the results add up to 18?

3. (from a CMU CS course handout\(^2\)) Suppose that we want to fill a basket with fruit, but we impose on ourselves some very quirky constraints:

\(^2\)http://www.cs.cmu.edu/afs/cs.cmu.edu/academic/class/15251-s09/Site/Materials/Handouts/generatingfunctions.pdf
(a) The number of apples must be a multiple of five (an apple a [week]day...)
(b) The number of bananas must be even (eaten before 15-251\(^3\) on Tues/Thurs...)
(c) We can take at most four oranges (too acidic...).
(d) There can be at most one pear (get mushy too fast...)

In how many ways can we do it?

4. Suppose the generating function for the sequence \(a_n\) is \(A(x)\), and for \(b_n\), it is \(B(x)\). What is the sequence behind the generating function \(A(x)B(x)\) (the product of the \(A(x)\) and \(B(x)\)) in terms of \(a_n\) and \(b_n\)? Multiplying two generating functions like this is called a ‘convolution’. Use convolutions to prove that

\[
\binom{r + s}{k} = \sum_{i=0}^{k} \binom{r}{i} \binom{s}{k - i}
\]

Also give a combinatorial proof of the relation above.

5. From Equation (3) and Equation (4), which sequence does the generating function \(\frac{1}{\sqrt{1-4x}}\) produce? Keeping in mind this result, in how many ways can you completely divide an \(n\)-gon into triangles by drawing non-intersecting diagonals\(^4\)?

Afterthoughts

Honestly, generating functions are a topic so vast that I have no hope of covering everything in my lifetime, but I hope I have inspired you with some of their magic. Many people do a vastly better job of explaining generating functions than I do, including:

- Qiaochu Yuan, the Olympiad math god and Math.SE superpower, in “Topics in generating functions”, which are notes from the Worldwide Online Olympiad Training (WOOT) program, available at [https://math.berkeley.edu/~qchu/TopicsInGF.pdf](https://math.berkeley.edu/~qchu/TopicsInGF.pdf)
- Luis von Ahn and Anupam Gupta of the famous CMU CS course 15-251, titled “Great Theoretical Ideas in Computer Science”, of which the notes on generating functions are available at [http://www.cs.cmu.edu/afs/cs.cmu.edu/academic/class/15251-s09/Site/Materials/Handouts/generatingfunctions.pdf](http://www.cs.cmu.edu/afs/cs.cmu.edu/academic/class/15251-s09/Site/Materials/Handouts/generatingfunctions.pdf)

(If you ever write a nice book, please do put it online.)

\(^3\)The CMU CS class. All the lecture slides, handouts, etc. from Spring 2009 are available online and it is a really cool course if you want to go through it.

\(^4\)You will have derived the Catalan numbers, which are immensely important in combinatorics, to the point where the book “Enumerative Combinatorics: Volume II” contains 66 combinatorial interpretations of the Catalan numbers.