Generating Functions Notes

Aaron Cohen (ac4810)

February 16, 2023

Definitions

Generating functions are a way to use our tools for dealing with functions to tackle combinatorics problems. Here is a definition:

> Definition: The generating function for the infinite sequence $\langle g_0, g_1, g_2, g_3 \dots \rangle$ is the power series:
>
> $$g(x) = g_0 + g_1 x + g_2 x^2 + g_3 x^3 + \dots$$

We will write $\langle g_0, g_1, g_2 \dots \rangle \leftrightarrow g(x)$ if they correspond in this way.

> Basic Examples:
>
> (1) $\langle 0, 0, 0, \dots \rangle \leftrightarrow 0 + 0x + 0x^2 + \dots = 0$
>
> (2) $\langle 1, 1, 1, \dots \rangle \leftrightarrow 1 + x + x^2 + \dots = \dfrac{1}{1-x}$ (geometric series)
>
> (3) $\langle 1, -1, 1, \dots \rangle = 1 - x + x^2 + \dots = 1 + (-x) + (-x)^2 + \dots = \dfrac{1}{1+x}$

We can combine simple sequences into larger ones using operations like function multiplication:

> Theorem 1 (Product rule): If $A(x)$ is a generating function for a sequence a_n, and $B(x)$ is a generating funciton for a sequence b_n, then $A(x)B(x)$ is a generating function for a new sequence c_n where:
>
> $$c_n = a_n b_0 + a_{n-1} b_1 + \dots a_0 b_n$$
>
> Proof:
>
> In polynomial multiplication, each term of $A(x)$ gets distributed to all terms of $B(x)$. A term with power x^n is produced exactly when the powers of the factors add up to n, i.e when the term $a_j x^j$ from $A(x)$ is multiplied by the term $b_{n-j} x^{n-j}$ from $B(x)$. Then in the resulting product, the x^n term will be the sum of all such terms. So $c_n = a_n b_0 + a_{n-1} b_1 + \dots a_0 b_n$.

This will allow us to create generating functions for a wide variety of problems. Let's do an example:

Application 1: Choosing from a set

Application 1 (Choosing from a set). Suppose we have a set with 3 elements: $\{a_0, a_1, a_2\}$. We can form a sequence where the nth term is the number of ways to choose n elements from the set. For a three element set, the sequence would be $\langle 1, 3, 3, 1, 0, 0 \dots \rangle$ because there is 1 way to choose 0 elements, 3 ways to choose 1 element, 3 ways to choose 2 elements, 1 way to choose 3 elements, and 0 ways to choose any number more than 3. We can use generating functions to generate this sequence for a set of any size k. We just need one result:

Theorem 2: If $A(x)$ represents choice from a set X, and $B(x)$ represents choice from a set Y, and $X \cap Y = \emptyset$ then $A(x)B(x)$ represents choice from $X \cup Y$.

Proof: To count the number of ways to choose n elements from $X \cup Y$, we can break up the problem into $n+1$ categories like this:

1. set of choices where all n are chosen from Y
2. set of choices where 1 element is from X and $n-1$ are from Y
3. set of choices where 2 are from X and $n-2$ are from Y

   $\vdots$

4. set of choices where all n are from X

The sum of the sizes of these sets is the number of ways to choose n from $X \cap Y$. We can calculate the size of each set very easily because we can simply multiply the number of ways to choose j from X by the number of ways to choose $n-j$ from Y. We already know those quantities because they are given by the jth and $n-j$th coefficients of $A(x)$ and $B(x)$ respectively. Therefore, the sum of the sizes of the sets, which is the number of ways to choose n from $X \cup Y$, is $a_n b_0 + a_{n-1} b_1 + \dots a_0 b_n$. Notice that this is the nth term of $A(x)B(x)$, as we proved in theorem 1. Therefore, $A(x)B(x)$ does indeed represent choice from the union of the sets.

Now we can easily tackle the original problem: for a set of size k, we can think of it as a union of k sets of size 1. Then the generating function representing choice from k elements can be constructed by multiplying k copies of the generating function representing choice from a 1 element set. The generating function representing choice from a 1-element set is $1+x$ because there is 1 way to choose 0 elements, and there is 1 way to choose 1. Therefore, choice from a set of size k is represented by the generating function $(1+x)^k$.

Note: if you know the binomial theorem, you know that when you expand out $(1+x)^k$ the coefficients are exactly $\binom{k}{n}$. So everything checks out.

Application 2: Choice with repetition

Now we are allowed to choose elements repeatedly. For example, for the set $\{a_0, a_1, a_2\}$ the sequence counting choice with repetition is $\langle 1, 3, 6, 10, \dots \rangle$. There are more choices now because, for example, choosing a_0 twice is now a valid way to choose 2 elements. However, we can still tackle this problem easily with the product rule:

Choice with repetition from a 1-element set $\leftrightarrow 1 + x + x^2 + \cdots = \dfrac{1}{1-x}$ since we can choose the single element any number of times. Then choice with repetition from a k element set is represented by $\dfrac{1}{(1-x)^k}$.

Unlike the previous example, it is not easy to see what sequence this generating function represents; we cannot simply expand terms. Instead, we can make use of taylor's theorem:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2} + \cdots + \frac{f^{(n)}(0)}{n!} + \ldots$$

So we can extract the coefficients from the generating function by taking the nth derivative at 0 and dividing by $n!$.

If you repeatedly take the derivative of $\dfrac{1}{(1-x)^k}$, you find a pattern that tells you that the nth coefficient is $\binom{n+k-1}{n}$. This same fact can be derived through the "stars and bars" argument, but we just did it from generating functions alone!

Partitions

In this section, we are going to use generating functions to count quantities related to partitions.

> Definition: A partition is a multiset (a set that can have multiple of the same element) of integers. A partition is said to be a partition of n if the sum of the elements of the partition is n.

In simple terms, a partition is a way of breaking up a number into smaller numbers. For example, (4,2,1) is a partition of 7.

Here are four quantities we will be able to count using generating functions:

1. $p(n)$: number of partitions of n
2. $p(n,k)$: number of partitions of n with k parts
3. $q(n)$: number of partitions of n with distinct parts
4. $q(n,k)$: number of partitions of n with k distinct parts

Counting $p(n)$

We will do this using a similar technique to how we did the binomial counting:

Claim: we can use the trick of multiplying single element sets to count partitions.

Proof sketch:
Suppose we have a generating function $A(x)$ representing the number of partitions that can be formed using only a subset X of the integers. For example, if X was {1,2} then the sequence would be $\langle 1, 1, 2, 2, 3, \ldots \rangle$ because there is one partition of 0, one partition of 1 (the partition (1)), two partitions of 2 ((1,1) and (2)), 2 partitions of 3 ((2, 1) and (1,1,1)), and 3 partitions of 4 ((1^4), (2,1,1), and (2,2)). And suppose $B(x)$ represents the sequence of the number of partitions that can be formed from a subset Y, which is disjoint from X. Then any partition using elements of $X \cup Y$ can be put into one of n sets: the sets where the elements from X contribute i and the elements from Y contribute $n - i$. Then just like before we find that $A(x)B(x)$ represents the sequence of the number of partitions that are formed from $X \cup Y$.

Essentially, we can use the same technique of breaking the set up into single elements! Now let's use it to count $p(n)$:

Given a single element set $\{i\}$, the sequence representing the number of partitions that can be formed from that set is $1 + x^i + x^{2i} + x^{3i} + \ldots$ because there is one way to make a partition of 0, one way to make a partition of i, one way to make a partition of $2i$, and so on. We can write this in closed form as $\frac{1}{x^i}$. Then we multiply all one-element sets to get the total number of partitions to see that $p(n)$ is represented by $\prod_{i=1}^{\infty} \frac{1}{1-x^i}$. Written in more rigorous notation, we say:

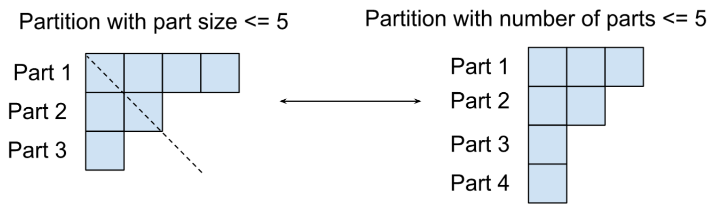
$$\sum p(n)x^n = \prod_{i=1}^{\infty} \frac{1}{1-x^i}$$

Counting $p(n, k)$

to calculate this quantity we will take a bit of a circuitous route. First, we will count partitions of n where the size of each part is no more than k. This is easy: we can simply multiply the generating functions for only the first k natural numbers:

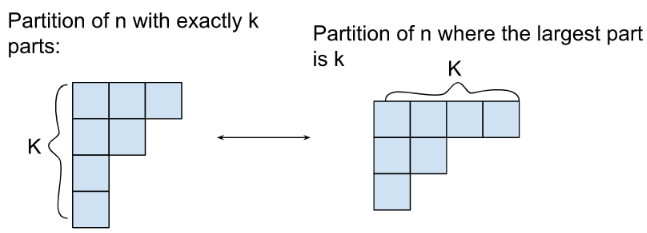
$$\prod_{i=1}^{k} \frac{1}{1-x^i}$$

Next, note that there is a bijection between the set of partitions of n where the parts are all at most k, and the set of partitions of n where the number of parts is at most k. This is illustrated in this diagram:



Where the map comes from flipping across the diagonal. Then our generating function that we just made actually counts this quantity too!

Next, note that flipping across the diagonal also gives us a bijection between partitions with exactly k parts (i.e $p(n, k)$), and partitions where the largest part is k:



Now, note that the number of partitions of n where the largest part is k is the same as the number of partitions of $n-k$ where all parts are at most k. This is because we can take any partition of $n-k$ where all parts are less than or equal to k, and then place a k part on top of it, and we will get a partition of n with largest part k. This is great, because we already know how to count partitions where all parts are less than or equal to k: it's $\prod_{i=1}^{k} \frac{1}{1-x^i}$. To form a generating function for $p(n, k)$, we just have to shift this generating function over by multiplying by x^k, so that the $(n-k)$th term of that function gets put in the nth place of this new one. So we have:

$$\sum p(n, k)x^n = x^k \prod_{i=1}^{k} \frac{1}{1-x^i}$$

Two-dimensional generating functions

We can also represent a two dimensional sequence like $p(n, k)$ with a two dimensional generating function. We can use x^n to keep track of the number we are partitioning, and use y^k to keep track of the number of parts. So, for example, if we expand out our generating function and find a term like $2x^4y^2$, that means that there are two partitions of 4 with two parts.

Let's do the same single-element trick again: given the set {i}, the corresponding generating function is $1x^0y^0 + 1x^iy^1 + 1x^{2i}y^2 \ldots$ because there is one way to make a partition of 0 with 0 parts, 1 way to mke a partition of i with one part, and 1 way to make a partition of 2i with two parts, etc. In closed form we have $\frac{1}{1-yx^i}$. Then taking the union we have:

$$\sum p(n,k)x^n y^k = \prod_{i=1}^{\infty} \frac{1}{1-yx^i}$$

Now we can relate our two expressions that we know expand out to the same sequence: we can use our one dimensional function to create a two dimensional function by using the whole sum as a coefficient for y^k:

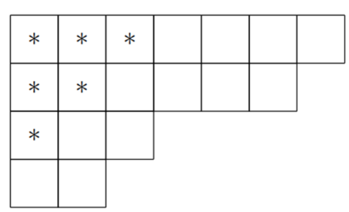
$$\sum p(n,k)x^n y^k = \sum_{k=1}^{\infty} y^k x^k \prod_{i=1}^{k} \frac{1}{1-x^i}$$

These two expressions are equivalent algebraically. We used combinatorics to prove an algebraic identity!

Distinct parts

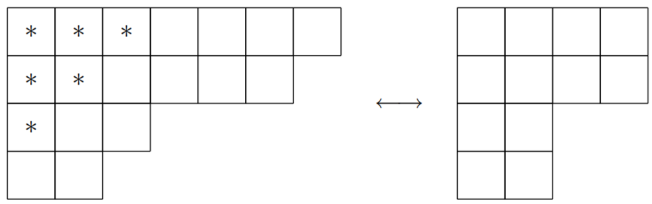
We can repeat what we just did except counting only partitions with distinct parts. We will end up getting another valid algebraic identity. Let's start by counting $q(n,k)$ directly. We will need two lemmas:

> Lemma 1: any partition with distinct parts contains a "staircase" inside it:



> Proof:
> The smallest part must be 1 or larger, and then every part after that must be larger than the previous by at least one. Therefore, the nth part is necessarily n or greater, meaning the staircase (which has length $n-1$ for the nth part) is contained within.

Lemma 2: any distinct partition minus the staircase gives you a valid (not necessarily distinct) partition with the same number of parts:



Proof:
We need to show that in the diagram resulting from deleting the staircase, the parts weakly increase from the bottom to top. This must be true since going up one level in the original distinct partition adds one to the size of the staircase and adds at least one to the length of the part. Therefore, the number of excess boxes outside of the staircase on each part can never decrease as we go up the partition. Hence, the resulting diagram is a valid partition.

Lemma 3: If you add a staircase to any partition, you get a distinct partition.

Proof:
Where there was an increase in part size (between adjacent parts) in the original partition, there is now a larger increase since the top has been boosted 1 more. Where there was previously two adjacent parts of the same size, there is now a gap of 1 between them. These are the only two cases, so we conclude that there is a difference between every adjacent part of the boosted partition therefore, it is distinct.

Combining these three lemmas, we conclude that there is a bijection between distinct partitions of n with k parts and arbitrary partitions with size $n - \binom{k}{2}$. We can then use our generating function for $p(n,k)$ and shift it by the size of the staircase to get:

$$\sum q(n,k)x^n = x^{k+\binom{k}{2}} \prod_{i=1}^{k} \frac{1}{1-x^i}$$

We can turn it two dimensional by summing with y^k again:

$$\sum q(n,k)x^n y^k = \sum_{k=1}^{\infty} y^k x^{k+\binom{k}{2}} \prod_{i=1}^{k} \frac{1}{1-x^i}$$

Now let's get the two dimensional version directly. {i} corresponds to $1 + x^i y^1$. Then we have:

$$\sum q(n,k)x^n y^k = \prod_{i=1}^{\infty} 1 + yx^i$$

So we have proven the following algebraic identity:

$$\prod_{i=1}^{\infty} 1 + yx^i = \sum_{k=1}^{\infty} y^k x^{k+\binom{k}{2}} \prod_{i=1}^{k} \frac{1}{1-x^i}$$

A surprising theorem

We just used combinatorics to prove an algebraic statement. Now we will use algebra to prove a combinatorics fact:

> Theorem 3: for all n, o(n), the number of partitions with all odd parts, is equal to q(n), the number of partitions with distinct parts
>
> Proof:
> Let's start by getting a generating function for $q(n)$. {i} corresponds to $1 + x^i$ because we can only use i once. Then:
>
> $$\sum q(n)x^n = \prod_{i=1}^{\infty}(1 + x^i)$$
>
> Now let's make a generating function for odd partitions. We can simply multiply all odd generating functions $\frac{1}{1-x^i}$, which gives us:
>
> $$\sum o(n)x^n = \frac{1}{(1-x)(1-x^3)(1-x^5)\cdots}$$
>
> Then we can multiply all the even terms to the numerator and denominator:
>
> $$\sum o(n)x^n = \frac{(1-x^2)(1-x^4)\cdots}{(1-x)(1-x^2)(1-x^3)(1-x^4)\cdots}$$
>
> Then, we can use difference of two squares on each term in the numerator, splitting each into $(1 - x^{i/2})(1 + x^{i/2})$. Then each $1 - x$ term will cancel out with an element in the denominator. This leaves us with only the $(1 + x)$ terms:
>
> $$\sum o(n)x^n = (1+x)(1+x^2)\cdots = \prod_{i=1}^{\infty} 1 + x^i$$
>
> This is exactly the same as our expression for $q(n)$, the number of odd partitions. Therefore, $q(n) = o(n)$ for all n !!!