

Notes on generating functions

My intention here is to just give you a brief introduction to pique your interest and familiarize you with the ideas. Check out the reference by Vitter and Flajolet [3] for further and more comprehensive info on generating functions. Another reference for discrete math in general is the text Concrete Math [2]. I expect you could find Mathematica or Maple packages to play with on the net, one possible lead is [1].

Example — the symbolic method. How many distinct n -node binary trees are there? The *ordinary generating function* for the set of binary trees is

$$b(z) = \sum_{n=0}^{\infty} b_n z^n,$$

where b_n is the number of binary trees of size n . In the function b , there is one term z^n of “size” n for each tree of size n . Postponing for a minute the question of what purpose this serves, can we find a simpler form for the function b ? Here is a simple set of rules defining the binary trees:

1. The tree consisting of a single node is a binary tree.
2. If a tree T is a binary tree, then the tree formed by creating a new root r and making T the single subtree of r is a binary tree.
3. If trees T_1 and T_2 are binary trees, then the tree formed by creating a new root r and making T_1 and T_2 the left and right subtrees of r is a binary tree.

From this set of rules we can see the following equivalence:

$$B \equiv \{\cdot\} \cup (\{\cdot\} \times B) \cup (\{\cdot\} \times B \times B).$$

Here “ \cdot ” represents a single node and “ \equiv ” means there is a size-preserving bijection (a one-to-one and onto function f such that the size of x equals the size of $f(x)$) between the set on the left and the set on the right. Replacing each root “ \cdot ” by a “ z ” (a term of “size” 1) and replacing the set union by addition yields the corresponding equivalence:

$$b(z) = z + z \times b(z) + z \times b(z) \times b(z).$$

Solving for $b(z)$ using the quadratic formula yields

$$b(z) = \frac{1 - z \pm \sqrt{1 - 2z - 3z^2}}{2z}.$$

From this closed form and a little algebra, we can see that $b(z)$ is well-defined as long as $z \leq 1/3$ — for larger values, the term inside the square root becomes negative. Amazingly, from this it follows directly (for reasons discussed below) that b_n grows roughly like 3^n . With a little more work, an exact form for b_n can be obtained.

Example — recurrence relations. The Fibonacci numbers are defined by the following recurrence:

$$f_n = \begin{cases} 1 & n = 0 \\ f_{n-1} + f_{n-2} & n > 0 \\ 0 & n < 0. \end{cases}$$

The corresponding generating function is $f(z) = \sum_n^\infty f_n z^n$. Can we find a simple form for $f(z)$? Using the recurrence, we get

$$\begin{aligned} f(z) &= \sum_n f_n z^n \\ &= f_0 + \sum_{n>0} (f_{n-1} + f_{n-2}) z^n \\ &= 1 + \sum_n f_{n-1} z^n + \sum_n f_{n-2} z^n \\ &= 1 + z f(z) + z^2 f(z). \end{aligned}$$

Solving for f yields

$$f(z) = \frac{1}{1 - z - z^2}.$$

From this we can see that f is well-defined as long as z is less than $r = \frac{\sqrt{5}-1}{2}$ — the smallest root of $1 - z - z^2$. From the general principle described below, it follows that f_n grows roughly like $(1/r)^n \approx 1.62^n$.

Estimating the rate of growth from the smallest singularity. In the above examples, we used the following rule of thumb:

Principle 1 (*Singularity principle*) *Let $f(z)$ be any generating function with non-negative coefficients.*

If $f(a)$ is well-defined for some constant a , then f_n is $o(a^{-n})$.

If $f(b)$ is not well-defined for some constant b , then f_n is not $O(1/c^n)$ for any constant $c > b$.

We can summarize this by saying that the best exponential approximation to the rate of growth of f_n is $1/r^n$, where r is the smallest singularity of $f(z)$, in other words, r is the largest value such that for all $z < r$, $f(z)$ is well-defined.

Why is this the case? If $f(a)$ is well-defined, then the infinite sum $\sum_n f_n a^n$ converges. Its terms must tend to zero as n tends to infinity:

$$\lim_{n \rightarrow \infty} f_n a^n = \lim_{n \rightarrow \infty} \frac{f_n}{a^{-n}} = 0.$$

This establishes the first part of the principle: $f_n = o(a^{-n})$.

On the other hand, let b and c be as in the second part of the principle: $f(b)$ is not well-defined and $b < c$. Suppose for contradiction that f_n is $O(c^{-n})$. Then $f_n b^n = O((b/c)^n)$, so that

$$\sum_n f_n b^n = O\left(\sum_n (b/c)^n\right) = O\left(\frac{1}{1 - b/c}\right) = O(1).$$

This contradicts our assumption that $f(b)$ was not well-defined.

Symbolic derivations — general principles. Here are a few general principles for *deriving* generating functions: Suppose sets A and B have generating functions $a(z)$ and $b(z)$, respectively.

- $A \cup B$ has generating function $a(z) + b(z)$, provided A and B are disjoint.
- $A \times B$ (the Cartesian product) has generating function $a(z)b(z)$.
- A^k (the Cartesian product of A with itself k times) has generating function $a(z)^k$.
- $A^* = \cup_{i \geq 0} A^i$ has generating function $\frac{1}{1 - a(z)}$. (A^* is called the Kleene closure of A , its elements are the finite sequences of elements from A)

(The rule for Cartesian products assumes that the “size” of an element (α, β) is the sum of the sizes of α and β . This was the case in the binary tree example. The subsequent rules have a similar assumption.)

Why are these principles true? Take the Cartesian product rule for example. For every element $\alpha \in A$ and $\beta \in B$, there is a term z^i in $a(z)$ and a term z^j in $b(z)$, where i and j are the sizes of α and β , respectively.

In the product $a(z)b(z)$, the pair (α, β) thus contributes a single term z^{i+j} of “size” equal to the size of the pair (α, β) . More formally, we can write

$$a(z)b(z) = \left(\sum_i a_i z^i \right) \left(\sum_j b_j z^j \right) = \sum_{i,j} a_i b_j z^{i+j} = \sum_n \left(\sum_{i+j=n} a_i b_j \right) z^n.$$

Since the number of pairs of size n in $A \times B$ is $\sum_{i+j=n} a_i b_j$, this proves the principle.

A canonical example. How many k -digit decimal numbers are there whose digits sum to n ? Let D be the set of digits $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$. For this problem, we think of each digit as having a size equal to itself. With this interpretation, the generating function for $d(z)$ is $1 + z + z^2 + \dots + z^9 = \frac{1-z^{10}}{1-z}$.

The set of k -digit numbers is then D^k . The generating function for D^k is $d(z)^k = \left(\frac{1-z^{10}}{1-z} \right)^k$. Here the “size” we associate with a k -digit number is just the sum of the sizes of the digits.

The singularity principle says that the n th coefficient grows like 1^n — this time it’s not so useful! On the other hand, note that for fixed k , this generating function has only finitely many terms, so we have to be a little more careful when asking questions about the asymptotics.

Arbitrary-degree rooted trees. Let t_n be the number of distinct rooted trees of size n . We can define the set of rooted trees as follows:

1. A single node is a rooted tree.
2. If T_1, T_2, \dots, T_k is any finite sequence of rooted trees, then a new rooted tree can be formed by creating a new root node r and making the roots of the trees in the sequence the children of r .

In fact, the first rule is the special case of the second that occurs when $k = 0$, so we can omit it. This gives the equivalence for the set T of rooted trees:

$$T \equiv \{\cdot\} \times T^*.$$

(Recall that T^* represents finite sequences of elements of T .) The generating function $t(z)$ thus satisfies

$$t(z) = z \times \frac{1}{1-t(z)}.$$

Solving for $t(z)$ using the quadratic formula yields

$$t(z) = \frac{1 \pm \sqrt{1-4z}}{2}.$$

From the singularity principle, it follows that t_n grows roughly like 4^n .

Exact answers. The singularity principle is useful for getting rough estimates of the exponential rate of growth of the coefficients of a generating function. Even more can be said about a generating function from its singularities, however, the proper tool for this is complex analysis and we don’t pursue it further here.¹

Instead we give a few simple examples where exact expressions can be obtained for the coefficients.

¹Some refinement is possible even without complex analysis. Here’s a hint about pinning down the polynomial term.

Suppose that $f(r)$ is well-defined but $f(z)$ is not well-defined for $z > r$. Compare the sum $\sum_n f_n/r^n$ (which converges) to the sum $\sum_n 1/n$ (which diverges). This will give you better upper bounds on f_n . Suppose further that f' (the derivative of f) is *not* well-defined at r . Compare the sum $\sum_n n f_n/r^{n-1}$ (the derivative at r , which diverges) to $\sum_n 1/n^2$ (which converges). This will give you better lower bounds.

If you are not so lucky that $f(r)$ is well-defined and $f'(r)$ isn’t, then consider integrating or differentiating f several times so that the resulting function has the property you want. Recall that each differentiation introduces a linear term in each coefficient, whereas integration factors one out.

A few standard series.

- $\frac{1}{1-z} = \sum_n z^n$. This is the generating function for any set where there is exactly one item of each size.
- $e^z = \sum_n z^n/n!$. This function is an example of an *exponential* generating function, which we don't go into here. These kinds of generating functions are useful for counting "labelled" structures — for instance, binary trees where each node is assigned a number; another example is counting restricted classes of functions or permutations.
- $(1+z)^m = \sum_n \binom{m}{n} z^n$. This is the generating function for the binomial coefficients. When m is an integer, the n th coefficient represents the number of size- n subsets of a set of size m .

Differentiation. Suppose you have a simple form $f(z)$ for the generating function $\sum_n f_n z^n$, and you want a simple form for $\sum_n n f_n z^n$ — that is, you want to introduce a linear term. Note that $f'(z) = \sum_n n f_n z^{n-1}$, so that the answer you want is just $z f'(z)$. For instance, applying this to the generating function $1/(1-z) = \sum_n z^n$ gives $z/(1-z)^2 = \sum_n n z^n$.

You can also use differentiation to get the exact form for the coefficients. The general rule is

Theorem 1 If $f(z) = \sum_n f_n z^n$, then

$$f_n = \frac{f^{(k)}(0)}{n!}.$$

Here $f^{(k)}$ represents the k th derivative of f . To prove it, first use induction to show that $f^{(k)}(z) = \sum_n n(n-1)\cdots(n-k+1)f_n z^{n-k}$, then substitute $k = n$ and $z = 0$ (so all terms but the leading one drop out) to get $f^{(k)}(0) = n!f_n$.

Repeated differentiation can often be messy, but not always. For example,

Theorem 2 For any m , the coefficient of z^n in $(1+z)^m$ is

$$\binom{m}{n} = \frac{m(m-1)\cdots(m-n+1)}{n!}.$$

This holds for arbitrary m , not just integer values. To prove it, just note that differentiating $(1+z)^m$ n times yields $m(m-1)\cdots(m-n+1)(1+z)^{m-n}$ and then apply the preceding theorem.

Substitution. From $\sum_n z^n = \frac{1}{1-z}$ it follows by substituting $2z$ for z that $\sum_n 2^n z^n = \frac{1}{1-2z}$. Similarly by substituting z^2 for z one obtains $\sum_n z^{2n} = \frac{1}{1-z^2}$.

Partial fractions. Recall that the generating function for the Fibonacci numbers is

$$f(z) = \frac{1}{1-z-z^2}.$$

Factor the denominator and separate into partial fractions:

$$\begin{aligned} f(z) &= \frac{1}{1-z-z^2} \\ &= \frac{1}{(1-az)(1-bz)} \\ &= \frac{a/(a-b)}{1-az} + \frac{b/(b-a)}{1-bz} \end{aligned}$$

Here $1/a$ and $1/b$ are the roots of $1-z-z^2$: $(1 \pm \sqrt{5})/2$. Letting $[z^n]f(z)$ denote the coefficient of z^n in $f(z)$, we have

$$\begin{aligned} [z^n]f(z) &= \frac{a/(a-b)}{1-az} + \frac{b/(b-a)}{1-bz} \\ &= \frac{a}{a-b}[z^n]\frac{1}{1-az} + \frac{b}{b-a}[z^n]\frac{1}{1-bz} \\ &= \frac{a}{a-b}a^n + \frac{b}{b-a}b^n \\ &= \frac{a^n - b^n}{a-b}. \end{aligned}$$

References

- [1] J. S. Devitt. Combinatorial objects and their generating functions: A Maple class room environment. In Thomas Lee, editor, *Mathematical Computation with Maple V: Ideas and Applications: Proceedings of the Maple Summer Workshop and Symposium, University of Michigan, Ann Arbor, June 28-30, 1993*, pages 20-26, Boston, MA, USA, 1993. Birkhäuser.
- [2] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading, MA, USA, second edition, 1994.
- [3] Jeffrey Scott Vitter and Philippe Flajolet. *Average-Case Analysis of Algorithms and Data Structures*, chapter 9. Elsevier Science Publishers B.V., 1990. ISBN 0-444-88075-5.