NOTES ON COMBINATORICS

LEO GOLDMAKHER

1. PASCAL'S TRIANGLE

Recall from high school (or before?) that $(x+y)^2 = x^2 + 2xy + y^2$. We can compute other powers, too:

$$
\begin{aligned}
(x+y)^0 &= 1 \\
(x+y)^1 &= x + y \\
(x+y)^2 &= x^2 + 2xy + y^2 \\
(x+y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3 \\
(x+y)^4 &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + x^4 \\
(x+y)^5 &= x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5
\end{aligned}
$$

It's clear what's happening to the variables – in the expansion of $(x+y)^n$, the exponent of x decreases from n to 0 as the exponent of y increases from 0 to n. But what are the coefficients? Erasing all the non-mysterious stuff, we're left with a triangle of numbers:

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
 & & & & & & \vdots & & & & & &
\end{array}
$$

In France (and most of the western world) this is called Pascal's triangle, because the first French person to study it was Levi ben Gershon (400 years before Pascal). In China this is called Yang Hui's triangle, because it was first investigated by Jia Xian. In Iran it's called Khayyam's triangle, because it was first investigated by Al-Karaji. But actually this triangle was already known in ancient Greece by the 2nd century BC, and in India by the 6th century AD.

In any event, there are tons of patterns in Pascal's triangle. Here are a few:

(1) Recursive property: each entry is the sum of two numbers that are above left and above right.
(2) The sum of all the numbers in the n^{th} row is 2^n.
(3) Each row is a palindrome.
(4) For any prime p, the entries in the p^{th} row (other than the 1's at each end) are all multiples of p.

A couple more patterns become apparent if we squash Pascal's triangle up against a wall; I call this the 'right Pascal's triangle':

$$\begin{array}{llllllll}
1 \\
1 & 1 \\
1 & 2 & 1 \\
1 & 3 & 3 & 1 \\
1 & 4 & 6 & 4 & 1 \\
1 & 5 & 10 & 10 & 5 & 1 \\
1 & 6 & 15 & 20 & 15 & 6 & 1 \\
1 & 7 & 21 & 35 & 35 & 21 & 7 & 1 \\
 & & & \vdots
\end{array}$$

Right Pascal's triangle

(5) The first column of the right Pascal's triangle consists of the counting numbers; the second column of the triangular numbers; the third column of the 'tetrahedral numbers'; etc.
(6) The sums along diagonals of the right Pascal's triangle produces the Fibonacci numbers.

Having observed a bunch of patterns, we can try to prove them. Let's take them in the same order as above:

(1) Recursive property: this is not clear how to prove.
(2) This follows from the definition of the numbers in Pascal's triangle: plugging the choices $x = y = 1$ into $(x + y)^n$, we see that $(1 + 1)^n$ is the sum of the n^{th} row of Pascal's triangle.
(3) This follows from the fact that $(x + y)^n = (y + x)^n$.
(4) This isn't clear how to prove.
(5) This follows from the recursive property.
(6) This isn't clear how to prove.

So, some of our observations aren't so deep, but others are less obvious. To prove them, we need to analyze where the coefficients of $(x + y)^n$ come from.

2. Refresher on polynomial multiplication

What do you get when you fully expand the product $(a_1 + a_2)(b_1 + b_2 + b_3)$? We can multiply the entire second factor by a_1, then multiply the entire second factor by a_2, and then add these to find

$$(a_1 + a_2)(b_1 + b_2 + b_3) = a_1b_1 + a_1b_2 + a_1b_3 + a_2b_1 + a_2b_2 + a_2b_3.$$

Awesome. What about a more complicated product?

$$(a_1 + a_2)(b_1 + b_2 + b_3)(c_1 + c_2) =$$
$$a_1b_1c_1 + a_1b_2c_1 + a_1b_3c_1 + a_2b_1c_1 + a_2b_2c_1 + a_2b_3c_1 + a_1b_1c_2 + a_1b_2c_2 + a_1b_3c_2 + a_2b_1c_2 + a_2b_2c_2 + a_2b_3c_2.$$

The main takeaway here is that

the product of a bunch of sums is the sum of all possible products of one term from each factor.

Thus, for example,

$$(x + y)^{17} = \underbrace{(x + y)(x + y) \cdots (x + y)}_{\text{17 factors}} = x^{17} + \_x^{16}y + \_x^{15}y^2 + \_x^{14}y^3 + \cdots + \_xy^{16} + y^{17},$$

where the coefficient of $x^k y^{17-k}$ is the number of ways of picking k x's out of the 17 factors. This begs for a notation:

Definition. The symbol $\binom{n}{k}$ denotes the number of ways of selecting k objects out of a group of n objects. (Read: n *choose* k.)

For example, what's $\binom{4}{2}$? For concreteness, let's say our 4 objects are the letters

M A T H

How many ways can we pick 2 of these? Six ways: MA, MT, MH, AT, AH, TH. Note that the order in which we pick the two letters isn't relevant – we're only interested in how many different subsets of size 2 there are of a given set of size 4.

With this notation in hand, we can now state

Theorem 2.1 (Newton's Binomial Theorem)**.**

$$(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y^1 + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n}y^n$$

This allows us to rewrite Pascal's triangle:

$$\begin{array}{ccccccccccccc}
 & & & & & & \binom{0}{0} & & & & & & \\
 & & & & & \binom{1}{0} & & \binom{1}{1} & & & & & \\
 & & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & & & \\
 & & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & & & \\
 & & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} & & \\
 & \binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} & \\
\binom{6}{0} & & \binom{6}{1} & & \binom{6}{2} & & \binom{6}{3} & & \binom{6}{4} & & \binom{6}{5} & & \binom{6}{6} \\
 & & & & & & \vdots & & & & & &
\end{array}$$

This looks like progress, but remember that $\binom{n}{k}$ is just a notation for a concept. Is there a nice way to actually compute $\binom{n}{k}$, other than simply counting all the possibilities by hand? As a warm-up, consider $\binom{8}{3}$. We will find a formula for it by answer the following question in two different ways:

Question. *How many ways are there to fill eight empty slots with eight given objects?*

To make this more precise, imagine 8 letters:

W H E N S D A Y

How many ways can we fill in the eight blanks

$$\_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_\ \ \_$$

with these letters?

<u>METHOD 1:</u> We put letters into each slot, one slot at a time. There are 8 choices for the first slot; 7 for the second; 6 for the third; etc. All told, this gives us $8 \cdot 7 \cdot 6 \cdots 2 \cdot 1 = 8!$ different ways to fill in the slots.

<u>METHOD 2:</u> We put letters into the first three slots, and then into the remaining five. To visualize this, split the 8 slots into two groups:

$$\_\ \ \_\ \ \_\ \Big|\ \_\ \ \_\ \ \_\ \ \_\ \ \_$$

How many ways are there to fill in the first three slots? Well, we first choose 3 elements out of the 8; by definition there are $\binom{8}{3}$ ways to do this. Once we've selected these three objects, how many ways are there of putting them into these three slots? $3! = 6$. Having filled in the first three slots, how many ways are there of filling in the last five slots using the five remaining letters? $5! = 120$. Note that we no longer have to choose these five elements – by choosing the first three we've locked ourselves into the last five. Putting this all together we get $\binom{8}{3} \cdot (3!) \cdot (5!)$ different ways to fill in the eight slots.

Of course, the outcomes of methods 1 and 2 have to agree, so we see that

$$\binom{8}{3} \cdot (3!) \cdot (5!) = 8!$$

We can run the same argument more generally: given n objects, there are $n!$ ways to fill in the slots

$$\underbrace{\_\ \_\ \_\ \cdots\ \_\ \_}_{n}$$

On the other hand, we can split up these n slots into two groups:

$$\underbrace{\_\ \_\ \cdots\ \_}_{k} \quad \Big| \quad \underbrace{\_\ \_\ \cdots\ \_}_{n-k}$$

To fill in the first k slots, we pick k objects out of our n – there are precisely $\binom{n}{k}$ ways of doing so – and then fill in the first k slots with these (there are precisely $k!$ ways of doing this). There are $(n-k)!$ ways of sorting the remaining $n-k$ objects into the remaining $n-k$ slots. Thus we get

$$\binom{n}{k} \cdot k! \cdot (n-k)! = n!,$$

which proves

Theorem 2.2. $\binom{n}{k} = \frac{n!}{k!(n-k)!}$

The strategy we used to prove this theorem is a classic example of a combinatorial argument: look at the same set in two different ways, and thus get two different expressions of counting the same objects. In this case, the set of objects we were counting was the number of ways to sort n things into n boxes. In our proof of Fermat's Little Theorem, we carried out a similar strategy by viewing the product of all the elements in the row of the multiplication table (mod p) in two different ways.

Remark. For a given k, the binomial coefficient $\binom{n}{k}$ can often be simplified. For example,

$$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)\cdot(n-2)!}{2\cdot(n-2)!} = \frac{n(n-1)}{2},$$

which you might recognize as the $(n-1)^{\text{st}}$ triangular number.

2.1. **Applications of binomial coefficients.** To demonstrate the power of the material we've developed thus far, we consider two quick examples of unexpected situations in which binomial coefficients are useful.

Example 1. (Counting subsets) How many subsets of $\mathcal{A}$ are there? We've seen that it's $2^{|\mathcal{A}|}$, but haven't proved this yet. The binomial theorem offers a very quick proof.

Theorem 2.3. *The number of subsets of $\mathcal{A}$ is $2^{|\mathcal{A}|}$.*

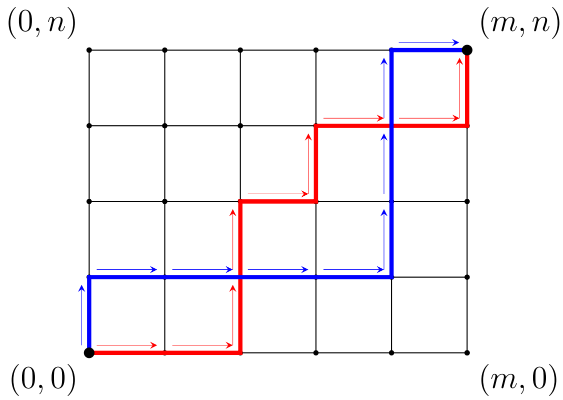
Proof. Let $n := |\mathcal{A}|$. There are $\binom{n}{k}$ subsets of $\mathcal{A}$ of size k. Thus, the total number of subsets of $\mathcal{A}$ is

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \cdots + \binom{n}{n}.$$

But taking $x = y = 1$ in Newton's binomial theorem, we see this sum is simply 2^n. This concludes the proof. $\square$

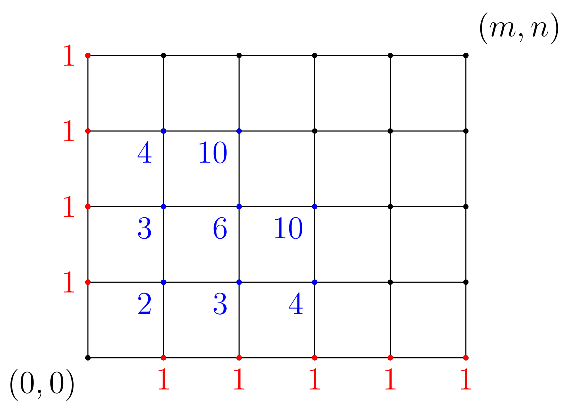
Example 2. (City walks) Suppose you're trying to walk from one intersection to another in a grid-designed city. Every street runs either east-west or north-south. How many different optimal routes are there? (A route is non-optimal if it takes longer than an alternative route.)

To make this question more precise, let's assume that the city is laid out in a regular grid, and that you're trying to get from $(0, 0)$ to (m, n). Here's an illustration of two different optimal routes:



One immediate observation is that any optimal route will only involve moving east and north. Thus we can rephrase the question: how many possible routes are there from $(0, 0)$ to (m, n) in which we're only allowed to walk east or north? We give two approaches to this problem.

APPROACH 1: RECURSION. Let's label each vertex in the figure with the number of ways of getting there from the origin (only allowing east or north steps). The vertices along the left and bottom edges (labeled in red in the diagram below) are straightforward. The interior vertices (labeled in blue) are a bit trickier.



This labelling is Pascalian: each label is the sum of the labels on the vertex to the south and the vertex to the west. (Can you explain why?) In any event, carrying out this recursion until we've filled in the entire diagram, we will eventually label the upper right vertex (m, n). This label is the number of optimal routes.

APPROACH 2: COMBINATORIAL REASONING. Overall, our journey consists of m blocks east and n block north, for a total of $m + n$ blocks. You can think of the route as a sequence of $m + n$ letters, each of which is either E (east) or N (north), e.g. $ENEENE$. Indeed, this is a bijection between strings consisting of m E's and n N's and optimal routes between $(0, 0)$ and (m, n), so all we have to do is count how many such strings there are. This is equivalent to choosing m positions out of $m + n$ total positions to place E's into, which can be done in $\binom{m+n}{m}$ ways. Thus, the number of optimal routes from the origin to (m, n) is precisely $\binom{m+n}{m}$.

DEPT OF MATHEMATICS & STATISTICS, WILLIAMS COLLEGE, WILLIAMSTOWN, MA, USA
E-mail address: leo.goldmakher@williams.edu