MTH6109: Combinatorics

Professor R. A. Wilson

Autumn Semester 2011

Lecture 1, 26/09/11

Combinatorics is the branch of mathematics that looks at *combining* objects according to certain rules (for example: sudoku).

Typically we want to decide *if* this can be done, or more generally, in *how many ways* it can be done.

Contents

Chapter 1

Counting sequences, subsets, integer partitions, and permutations

1.1 Sequences

Example 1 *How many sequences (order is important) of length* 5 *can we make using letters of the alphabet?*

Answer: Since repetition is allowed, each position can be filled by one of the 26 letters, hence the answer is 26^5.

Example 2 *How many sequences in Example 1 use each letter at most once?*

Answer: $26 \times 25 \times 24 \times 23 \times 22$.

MULTIPLICATION PRINCIPLE: We want to count the number of elements of some set X. Suppose we can generate the elements of X using a process consisting of n stages such that:

(a) The number of choices in the ith stage is t_i, and this number is independent of choices we make in the previous stages

(b) If we make a different choice at any stage we get a different element of X.

Then $|X| = t_1.t_2.\cdots t_n$.

Example 3 *How many sequences in Example 2 start and end with a vowel?*

(Wrong) answer: $5 \times 25 \times 24 \times 23 \times 4$. This is wrong because we could have chosen a vowel for one of the middle letters, and then there would be fewer than 4 choices for the last letter.

Correct answer:

Stage 1: Choose 1st letter: 5 choices.

Stage 2: Choose 5th letter: 4 choices.

Stage 3: Choose 2nd letter: $24 = 26 - 2$ choices.

Stage 4: Choose 3rd letter: 23 choices.

Stage 5: Choose 4th letter: 22 choices.

Answer: $5 \times 4 \times 24 \times 23 \times 22$.

Theorem 1 *Let S be a set with m elements. Then the number of ordered sequences of length k using the elements of S is*

(a) m^k if we can use each element an arbitrary number of times.

(b) $m.(m-1).\cdots.(m-k+1)$ if we can use each element at most once.

Example 4 *How many sequences of length 5 using $\{A, B, \ldots, , Z\}$ are there, using each letter at most once, and containing exactly one vowel?*

Reasoning:

Stage 1: Choose position for the vowel: 5 choices.

Stage 2: Choose the vowel: 5 choices.

Stage 3: Choose a consonant for the first available position: 21 choices.

Stage 4: Choose a consonant for the next available position: 20 choices.

Stage 5: Choose a consonant for the next available position: 19 choices.

Stage 6: Choose a consonant for the next available position: 18 choices.

Answer: $5 \times 5 \times 21 \times 20 \times 19 \times 18$.

Example 5 *How many sequences of length 5 use each letter at most once and contain* at least *one vowel?*

(Wrong) answer:

Stage 1: Choose position for the vowel: 5 choices.

Stage 2: Choose the vowel: 5 choices.

Stage 3: Choose any other letter for the first available position: 25 choices.

Stage 4: Choose a letter for the next available position: 24 choices.

Stage 5: Choose a letter for the next available position: 23 choices.

Stage 6: Choose a letter for the next available position: 22 choices.

Answer: $5 \times 5 \times 25 \times 24 \times 23 \times 22$. Why is this wrong? Because it fails the second condition for the multiplication principle: we can get the same sequence in two different ways. For example, the sequence BEIOZ could be obtained by first choosing the third position to put a vowel, then choosing the vowel I, and then choosing the other letters. But it could also be obtained by first choosing the fourth position to put a vowel, then choosing the vowel O, and then choosing the other letters. So some sequences are being counted more than once.

Correct answer: (Total number of sequences with no repetition)
− (number of sequences with no vowels).

$$= (26 \times 25 \times 24 \times 23 \times 22) - (21 \times 20 \times 19 \times 18 \times 17).$$

Example 6 *How many sequences of length* 5 *use each letter at most once and contain at least two vowels?*

Answer: (Total number of sequences with no repetition)
− (number of sequences with no vowels)
− (number of sequences with exactly one vowel).

$$= (26 \times 25 \times 24 \times 23 \times 22) - (21 \times 20 \times 19 \times 18 \times 17) - (5 \times 5 \times 21 \times 20 \times 19 \times 18).$$

Definition 1 *Let X be a finite set. A* partition *of X is a family (i.e. set) of non-empty subsets $X_1, X_2, \ldots, X_n$ of X, such that every element of X belongs to exactly one of these subsets X_i.*

ADDITION PRINCIPLE: Suppose that X is a set and $\mathcal{F} = \{X_1, X_2, \ldots, X_n\}$ is a partition of X. Then

$$|X| = |X_1| + |X_2| + \cdots + |X_n|.$$

Example 6 revisited. Let X be the set of all sequences of length 5 with no repetitions. Let

- X_0 be the subset of sequences which have no vowels;
- X_1 be the subset of sequences which have exactly one vowel;
- X_2 be the subset of sequences which have at least two vowels.

Then $\{X_0, X_1, X_2\}$ is a partition of X, so by the addition principle

$$|X| = |X_0| + |X_1| + |X_2|.$$

Therefore $|X_2| = |X| - |X_0| - |X_1|$, and we have already calculated all the terms on the right-hand side, so we can calculate $|X_2|$.

Lecture 2, 27/09/11

CONCLUSION: When using the multiplication principle and/or the addition principle to count things, it is important to make sure that

(a) every element of the set has been counted, and

(b) no element has been counted twice.

Example 7 *If A is a set of k elements, and B is a set of m elements, how many functions are there from A to B?*

Answer: Write $A = \{a_1, a_2, \ldots, a_k\}$, and choose our function f by a k-stage process:

Stage 1: choose $f(a_1)$, in one of m ways;

Stage 2: choose $f(a_2)$, in one of m ways;

...

Stage k: choose $f(a_k)$, in one of m ways.

So the answer is m^k.

Example 8 *How many of the functions in Example 7 are injective?*

Stage 1: choose $f(a_1)$, in one of m ways;

Stage 2: choose $f(a_2)$, in one of $m-1$ ways;

...

Stage k: choose $f(a_k)$, in one of $m-k+1$ ways.

So the answer is $m.(m-1).\cdots.(m-k+1)$.

1.2 Subsets

Example 9 *Let X be a set with n elements. How many different subsets does X have?*

Answer: using the multiplication principle. Let $X = \{x_1, x_2, \ldots, x_n\}$. Then we choose our subset S by an n-stage process, as follows.

Stage 1: either $x_1 \in S$ or $x_1 \notin S$ (2 choices);

Stage 2: either $x_2 \in S$ or $x_2 \notin S$ (2 choices);

...

Stage n: either $x_n \in S$ or $x_n \notin S$ (2 choices).

Hence there are 2^n subsets altogether.

Another method: we make each subset correspond to a sequence of 0s and 1s as follows. The ith digit in the sequence is 1 if $x_i \in S$ and is 0 if $x_i \notin S$. (For example, if $n = 5$ and $S = \{x_1, x_4, x_5\}$, then the corresponding sequence is 10011.) Now it is easy to see that each subset gives a sequence, and given the sequence, we can work out the corresponding subset. In other words, we have a *bijection* between the set of subsets of X and the set of binary sequences of length n. Since the number of sequences is 2^n, it follows that the number of subsets of X is also 2^n.

CORRESPONDENCE PRINCIPLE: Let X and Y be finite sets. If there is a bijection $f : X \to Y$, then $|X| = |Y|$.

Theorem 2 *Let X be a set with n elements. Then the number of distinct subsets of X is 2^n.*

Proof. Let $\mathcal{P}$ be the set of all subsets of X. We want to show that $|\mathcal{P}| = 2^n$.

Let Y be the set of all sequences of length n using the symbols 0 and 1. We know from the previous section that $|Y| = 2^n$. We now show that $|\mathcal{P}| = |Y|$ by defining a bijection $f : \mathcal{P} \to Y$, as follows.

Let $X = \{x_1, x_2, \ldots, x_n\}$, and for each $S \subseteq X$ define $f(S) \in Y$ by putting $f(S) = b_1 b_2 \ldots b_n$, where $b_i = 1$ if $x_i \in S$ and $b_i = 0$ otherwise. Then $f : \mathcal{P} \to Y$ is a bijection (since it has an inverse g given by $g(b_1 b_2 \ldots b_n) = \{x_i \mid b_i = 1\}$).

Therefore, by the correspondence principle, $|\mathcal{P}| = |Y| = 2^n$.

Definition 2 *If n is a positive integer, we define 'n factorial' (written $n!$) by*

$$n! = n.(n-1).(n-2).\cdots.2.1,$$

that is, the product of all the positive integers up to and including n. We also define $0! = 1$.

Note that the number of sequences of length k (without repetition) from a set of size m is $m(m-1)\cdots(m-k+1) = m!/(m-k)!$.

Example 10 *How many subsets (order not important) of $\{A, B, C, \ldots, Z\}$ are there of size 5?*

Answer: We already know that number of *sequences* of length 5 (without repetition) is $26 \times 25 \times 24 \times 23 \times 22$.

But we could also count these sequences in a different way: first pick the set of 5 letters that are going to be used, say $\{l_1, l_2, l_3, l_4, l_5\}$, and then put them in order. Now each set can be ordered in 5! ways, and different subsets give rise to different sequences, so, by the multiplication principle:

$$\text{Number of sequences} = 5!(\text{Number of subsets})$$

and therefore

$$\text{Number of subsets} = (\text{Number of sequences})/5! = \frac{26 \times 25 \times 24 \times 23 \times 22}{5 \times 4 \times 3 \times 2 \times 1}.$$

This can also be written as

$$\frac{26!}{21!5!}.$$

Lecture 3, 29/9/2011

Theorem 3 *Let X be a set with m elements. Then for each k with $0 \leq k \leq m$, the number of subsets of X of size k is*

$$\frac{m!}{(m-k)!k!}.$$

Definition 3 *We write*

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

and read 'm choose k'.

Proof: essentially the same as the example. We count in two ways the number of (ordered) *sequences* of length k (without repetition). First, by direct counting (as above) the number is $m(m-1).\cdots.(m-k+1)$. Second, this equals

$$(\text{the number of subsets of size } k).k!$$

since there are $k!$ orderings of the elements $x_1, x_2, \ldots, x_k$ of the subset.

Hence the number of subsets of size k equals

$$\frac{m!}{k!(m-k)!}.$$

Another way of thinking of this is that we choose the 'first' element of the set in m ways, the 'second' in $m-1$ ways, and so on. But then we have counted each subset of size k exactly $k!$ times, so we must divide by $k!$.

Example 11 *How many sequences of length six can be made from the letters of FLEECE, using each letter exactly once?*

Answer: $6!/3! = 6 \times 5 \times 4 = 120$.

Reason 1: Choose the letters one at a time, in 6! ways. But each sequence is then counted 3! times, as if I re-order the three Es, the result is the same. Hence the answer is 6!/3!

Reason 2: Another way to count the arrangements is:

Stage 1: Choose the position to put the F (6 choices)

Stage 2: Choose the position to put the L (5 choices)

Stage 3: Choose the position to put the C (4 choices)

Then the remaining three letters are E, so the total number of choices is $6 \times 5 \times 4$.

Reason 3: Using the theorem, we count as follows:

Stage 1: Choose the three positions to put the Es (number of choices is $\binom{6}{3}$).

Stage 2: Choose the order of the other three letters in 3! ways.

Hence the answer is

$$\binom{6}{3} 3! = \frac{6!}{3!}.$$

Theorem 4 *(One form of the binomial theorem.) If m is any positive integer, then*

$$(1+x)^m = \sum_{k=0}^{m} \binom{m}{k} x^k.$$

Proof: Consider $(1+x)^m$ as a product of m factors $B_1.B_2.\cdots.B_m$, where

$$B_1 = B_2 = \cdots = B_n = (1+x).$$

To get a term x^k in this product we need to choose an x from exactly k of the factors $B_1, B_2, \ldots, B_m$, and a 1 from the remaining factors. The number of ways of doing this is

$$\binom{n}{k}.$$

Hence in the expansion of the product there are exactly $\binom{n}{k}$ terms x^k. In other words, the coefficient of x^k is $\binom{n}{k}$.

Corollary 1 *Put $x = 1$ to get*

$$2^m = \sum_{k=0}^{m} \binom{m}{k}.$$

A combinatorial interpretation of this result is that the left-hand side is the number of subsets of a set of size m, whereas the right-hand side is the sum over all possible sizes k of subsets, of the number of subsets of that size. So we don't really need the binomial theorem to prove this result: it is combinatorially obvious!

Corollary 2 *What happens if we put $x = -1$? We get*

$$0 = \binom{m}{0} - \binom{m}{1} + \binom{m}{2} - \binom{m}{3} + \cdots$$

which we can re-arrange as

$$\binom{m}{1} + \binom{m}{3} + \binom{m}{5} + \cdots = \binom{m}{0} + \binom{m}{2} + \binom{m}{4} + \cdots$$

that is the number of subsets of odd size is equal to the number of subsets of even size.

Lecture 4, 03/10/11

In fact, for m odd this is more or less obvious, but not so for m even.

Example 12 *If $m = 5$, we have $\binom{5}{1} + \binom{5}{3} + \binom{5}{5} = 5+10+1$ and $\binom{5}{0} + \binom{5}{2} + \binom{5}{4} = 1+10+5$, and we get not only equality of the sums, but equality of the individual terms, in reverse order.*

If $m = 6$, we have $\binom{6}{1} + \binom{6}{3} + \binom{6}{5} = 6 + 20 + 6$ and $\binom{6}{0} + \binom{6}{2} + \binom{6}{4} + \binom{6}{6} = 1 + 15 + 15 + 1$, and the equality is much less obvious.

Indeed, we have the following *binomial identity*:

$$\binom{n}{k} = \binom{n}{n-k}.$$

One way to see this is to set up a bijection f between the set A_k of subsets of size k and the set A_{n-k} of subsets of size $n-k$, by defining $f(Y) = Y^c$, the complement of Y. Clearly, if $|Y| = k$ then $|Y^c| = n-k$, and $(Y^c)^c = Y$, so this is a bijection. But $|A_k| = \binom{n}{k}$ and $|A_{n-k}| = \binom{n}{n-k}$, so the result follows.

An alternative method is to substitute into the formula:

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

so

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

Example 13 *Suppose I want to pick a team of k players from a class of m people, and a captain of the team. How many ways can this be done?*

One way to do this is pick the team first, which can be done in $\binom{m}{k}$ ways, since order is not important. Then we can choose any of the k members of the team to be the captain. Thus the answer is $k\binom{m}{k}$.

But this isn't necessarily the way you'd actually do it in practice: you might choose the captain first (m possible choices), and then choose the remaining $k-1$ members of the team from the other $m-1$ members of the class. hence the answer is $m\binom{m-1}{k-1}$.

Since we have counted the same things in two different ways, we must get the same answer, in other words

$$k\binom{m}{k} = m\binom{m-1}{k-1}.$$

If you're not convinced by this combinatorial argument, you can verify the identity using the formula:

$$k\binom{m}{k} = \frac{k.m!}{k!(m-k)!} = \frac{m!}{(k-1)!(m-k)!}$$

while

$$m\binom{m-1}{k-1} = \frac{m.(m-1)!}{(k-1)!((m-1)-(k-1))!} = \frac{m!}{(k-1)!(m-k)!}.$$

Example 14 *Suppose I want to pick a team of k people from a group of $m+1$ people, including myself. Either I include myself in the team or I do not. If I am in the team, there are $\binom{m}{k-1}$ ways of choosing the rest of the team. If I am not in the team, there are $\binom{m}{k}$ ways of choosing the whole team. Since exactly one of these two cases must occur, the total number of ways of choosing the team is*

$$\binom{m}{k-1} + \binom{m}{k}.$$

But we already know that the number of ways of choosing a team of k from $m+1$ people is $\binom{m+1}{k}$. Hence

$$\binom{m}{k-1} + \binom{m}{k} = \binom{m+1}{k}.$$

Again you can verify this using the formula if you prefer: the left-hand side comes to

$$\frac{m!}{(k-1)!(m-k+1)!} + \frac{m!}{k!(m-k)!} = \frac{m!}{k!(m-k+1)!}(k+(m-k+1)) = \frac{(m+1).m!}{k!(m+1-k)!}.$$

1.3 Counting ordered partitions of integers

Example 15 *How many solutions are there to the equation*

$$x_1 + x_2 + x_3 + x_4 = 11$$

where x_1, x_2, x_3, x_4 are non-negative integers?

Answer: given a solution, we draw a picture of 14 boxes in a row, of which 11 are empty and 3 contain plus signs. There are x_1 boxes before the first plus sign, x_2 between the first and second, x_3 between the second and third, and x_4 boxes after the last plus sign. Conversely, given such a picture, we get a solution to the equation, as we read off x_1 as the number of empty boxes before the first plus sign, etc. Hence each x_i is a non-negative integer, and since the total number of empty boxes is 11 we have $x_1 + x_2 + x_3 + x_4 = 11$.

Thus we have set up a bijection between the set of solutions to the equation, and the set of pictures. We can easily count the pictures, as all we need to do is choose which 3 of the 14 boxes have plus signs in them. So there are $\binom{14}{3}$ pictures. Hence (by the correspondence principle) there are $\binom{14}{3}$ solutions to the equation.

Lecture 5, 04/10/11

The same argument proves:

Theorem 5 *The number of solutions to the equation*

$$x_1 + x_2 + \cdots + x_k = n$$

where $x_1, x_2, \ldots, x_k$ are non-negative integers, and where n is a non-negative integer and k is a positive integer, is

$$\binom{n+k-1}{k-1}.$$

Example 16 *How many ways are there of distributing* 15 *(identical) bottles of beer among* 6 *(different) students?*

Answer: if the first student gets x_1 bottles, and so on, then we want the number of solutions in non-negative integers to the equation

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 15.$$

So the answer is

$$\binom{15+6-1}{6-1} = \binom{20}{5}.$$

Example 17 *In how many of these cases does every student get at least one bottle?*

Answer: let us give each student one bottle first, and then distribute the remaining 9 bottles in $\binom{9+6-1}{6-1} = \binom{14}{5}$ ways.

More generally, the number of solutions in *positive* integers to the equation

$$x_1 + x_2 + \cdots + x_k = n$$

is the same as the number of solutions in non-negative integers to the equation

$$y_1 + y_2 + \cdots + y_k = n - k,$$

where $y_i = x_i - 1$, that is

$$\binom{n-k+k-1}{k-1} = \binom{n-1}{k-1}.$$

1.4 Counting permutations

Definition 4 *A permutation of a set X is a bijection $f : X \to X$.*

Theorem 6 *If X is a set with n elements, then the number of permutations of X is $n!$.*

Proof: Let $X = \{x_1, x_2, \ldots, x_n\}$. Then the permutations of X can be chosen by an n-stage process:

Stage 1: Choose $f(x_1)$: there are n choices;

Stage 2: Choose $f(x_2)$: there are $n-1$ choices, since $f(x_2) \neq f(x_1)$;

Stage 3: Choose $f(x_3)$: there are $n-2$ choices;

...

Stage n: Choose $f(x_n)$: there is only 1 possibility left.

Example 18 *Let $X = \{1,2,3,4,5,6\}$, and define $f : X \to X$ by $f(1) = 1$, $f(2) = 4$, $f(3) = 5$, $f(4) = 6$, $f(5) = 3$, $f(6) = 2$. One way to write this is as*

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 4 & 5 & 6 & 3 & 2 \end{pmatrix}$$

where the top row just list the elements x of the set, and the bottom row the corresponding $f(x)$.

Another way is to draw a picture with dots for the six elements of the set, and arrows from x to $f(x)$ for each x. Then you see that the picture consists of a collection of cycles. *In this case the cycles are $(2,4,6)$ and $(3,5)$ and (1), so we write this permutation as $(2,4,6)(3,5)(1)$ or $(2,4,6)(3,5)$ for short.*

But note that we can start a cycle anywhere: $(2,4,6) = (4,6,2) = (6,2,4)$.

Definition 5 *A* cyclic permutation *of length k on a set X of size n is any permutation of the form $(x_1, x_2, \ldots, x_k)$, for some distinct $x_i \in X$, that is $f(x_i) = x_{i+1}$ for $i < k$, and $f(x_k) = x_1$.*

Example 19 *How many cyclic permutations of length 6 are there on the set $\{1,2,3,4,5,6\}$?*

Answer: each cycle can be written as $(x_1, x_2, x_3, x_4, x_5, x_6)$, for some ordering of the elements of the set. There are 6! such orderings. But each cycle is counted 6 times, so the number of cycles is $6!/6 = 5! = 120$.

Example 20 *How many permutations of $\{1,2,3,4,5,6\}$ are there of the shape $(x_1, x_2, x_3, x_4)(x_5, x_6)$?*

Answer: Again there are 6! ways of ordering the elements of the set. Now each permutation is counted 8 times, since the cycle of length 4 can start with any of the 4 elements x_1, x_2, x_3, x_4, and the cycle of length 2 can start with either x_5 or x_6.

Lecture 6, 06/10/11

Example 21 *How many permutations of the set $\{1,2,3,4,5,6\}$ can be expressed as the product of two disjoint cycles of length 3?*

Answer: Any such permutation is of the form $(x_1, x_2, x_3)(x_4, x_5, x_6)$, and we can choose the (ordered) sequence $x_1, x_2, x_3, x_4, x_5, x_6$ in 6! ways, as before.

But the cycle $(x_1, x_2, x_3) = (x_2, x_3, x_1) = (x_3, x_1, x_2)$ has three different names, and similarly the cycle (x_4, x_5, x_6) has three different names. This suggests the answer is $6!/9 = 80$. Is this right?

No! Because we also have

$$(x_1, x_2, x_3)(x_4, x_5, x_6) = (x_4, x_5, x_6)(x_1, x_2, x_3).$$

In other words, since the two cycles have the same length, we can interchange the two cycles, and hence get twice as many ways to represent the same permutation.

Thus (by the multiplication principle) each permutation can be written in $2 \times 3 \times 3$ ways: first choose which of the two cycles to put first, then choose which of the three possible numbers to start the first cycle with, then choose which of the three possible numbers to start the second cycle with.

Hence the answer is $6!/18 = 40$.

Chapter 2

Recurrence relations and generating functions

2.1 Examples of recurrence relations

Example 22 *(Fibonacci numbers) How many (ordered) sequences of* $1s$ *and* $2s$ *are there which sum to* n*, for a fixed integer* n*?*

Let this number be F_n. For small n we have:

n	Sequences summing to n	F_n
0	empty sequence	1
1	1	1
2	$1+1, 2$	2
3	$1+1+1, 1+2, 2+1$	3
4	$1+1+1+1, 1+1+2, 1+2+1, 2+1+1, 2+2$	5

Getting a formula for F_n directly is not so easy, but we can get a *recurrence relation* expressing F_n in terms of F_k for smaller k.

Claim: $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$.

Proof: Let S_n be the set of all (ordered) sequences of 1s and 2s summing to n, so that $F_n = |S_n|$.

Example: $S_3 = \{(1,1,1), (1,2), (2,1)\}$ and $F_3 = |S_3| = 3$.

Let X_n be the set of all sequences in S_n which end with a 1, and let Y_n be the set of all sequences in S_n which end with a 2, so that by the addition principle

$$|S_n| = |X_n| + |Y_n|.$$

Example: $X_3 = \{(1,1,1), (2,1)\}$ and $Y_3 = \{(1,2)\}$, and $|S_3| = 2 + 1 = 3$.

Now if $(x_1, x_2, \ldots, x_k) \in X_n$ then $x_k = 1$ and $x_1 + x_2 + \cdots + x_k = n$, so $x_1 + x_2 + \cdots + x_{k-1} = n - 1$. Hence the map $(x_1, x_2, \ldots, x_k) \mapsto (x_1, x_2, \ldots, x_{k-1})$

which deletes the last term of each sequence is a map from X_n to S_{n-1}. Is it obvious that this is a bijection? If not, how do we prove it?

So, by the correspondence principle, $|X_n| = |S_{n-1}|$.

Example: $g : X_3 \to S_2$ is the map given by $g(1,1,1) = (1,1)$, $g(2,1) = (2)$: just delete the last element in the sequence.

Similarly, the map $(x_1, x_2, \ldots, x_k) \mapsto (x_1, x_2, \ldots, x_{k-1})$ is a bijection from Y_n to S_{n-2}. Hence $|Y_n| = |S_{n-2}|$.

Putting all this together we get

$$F_n = |S_n| = |X_n| + |Y_n| = |S_{n-1}| + |S_{n-2}| = F_{n-1} + F_{n-2}.$$

Example 23 *Let G_n be the number of sequences of positive integers (of any length) which sum to n.*

Applying the same ideas as above, the last number in the sequence can be any of the numbers $1, 2, 3, \ldots, n$, and in these cases the rest of the sequence sums to $n-1, n-2, n-3, \ldots, 0$. Hence by the addition principle we have

$$G_n = G_{n-1} + G_{n-2} + \cdots + G_1 + G_0.$$

For small n, we calculate directly that $G_0 = 1$, $G_1 = 1$, $G_2 = 2$. Now use the recurrence relation to get $G_3 = 2 + 1 + 1 = 4$, $G_4 = 4 + 2 + 1 + 1 = 8$, etc. Do you see the pattern?

$G_n = 2^{n-1}$. Can you prove it? By induction on n?

Can you simplify the recurrence relation? Consider the expression for G_{n-1}:

$$G_{n-1} = G_{n-2} + G_{n-3} + \cdots + G_1 + G_0.$$

So all the terms from G_{n-2} onwards in the expression for G_n can be replaced by G_{n-1}, and we get

$$G_n = 2G_{n-1}.$$

Example 24 *We have already seen an example of a recurrence relation for binomial coefficients:*

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

A recurrence relation always gives you a recursive method of calculating the answer you want: in this case, you construct *Pascal's triangle* one row at a time. The nth row contains the numbers

$$\binom{n}{0}, \binom{n}{1}, \binom{n}{2}, \ldots, \binom{n}{n}$$

in order, and the recurrence relation tells you to calculate each entry as the sum of two specified entries in the row above.

A recurrence relation may or may not give you a nice formula for calculating the answer directly. In this case we have *used* the formula

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

to *deduce* the recurrence relation, but you could (and in general it is more useful to) use the *recurrence relation* to prove (by induction on n) the above formula.

Lecture 7, 10/10/11

Bell numbers: counting partitions of a set. The nth Bell number B_n is defined as the number of partitions of a set with n elements. The names of the elements do not matter, so we might as well suppose our set is $S_n = \{1, 2, \ldots, n\}$.

If $n = 0$, then $S_0 = \emptyset$, and there is a unique partition of S_0, namely $\emptyset$. Hence $B_0 = 1$.

If $n = 1$, then $S_1 = \{1\}$, and there is a unique partition of S_1, namely $\{\{1\}\}$, so $B_1 = 1$.

If $n = 2$, then $S_2 = \{1, 2\}$, and there are exactly two partitions of S_2, one partition into two pieces and one partition into just one part, that is $\{\{1\}, \{2\}\}$ and $\{\{1, 2\}\}$. So $B_2 = 2$.

When $n = 3$, we have one partition into a single part, $\{\{1, 2, 3\}\}$, and one partition into three parts, $\{\{1\}, \{2\}, \{3\}\}$, and three partitions into two parts, $\{\{1\}, \{2, 3\}\}$, $\{\{2\}, \{1, 3\}\}$, and $\{\{3\}, \{1, 2\}\}$. Therefore $B_3 = 5$.

Theorem 7 *The Bell numbers satisfy the following recurrence relation:*

$$B_n = \sum_{k=1}^{n} \binom{n-1}{k-1} B_{n-k}.$$

Proof: Let $S_n = \{1, 2, 3 \ldots, n\}$ and let P_n be the set of all partitions of S_n, so that $B_n = |P_n|$.

Now we partition P_n according to the size of the part of the partition containing n: let T_k be the set of those partitions $\mathcal{F}$ of S_n which have the property that the part of $\mathcal{F}$ which contains n has size k. In symbols,

$$T_k = \{\mathcal{F} \mid n \in X \in \mathcal{F} \text{ with } |X| = k\}.$$

Now *every* partition $\mathcal{F}$ of S_n has a unique part $X \in \mathcal{F}$ such that $n \in X$, and this part must have *some* size, between 1 and n inclusive. So

$$B_n = |P_n| = \sum_{k=1}^{n} |T_k|.$$

Next we need to work out the size of T_k, for each k. We can pick the partitions in T_k by a two-stage process: first pick the part of the partition which contains n; then pick the rest of the partition.

Stage 1: We need to pick the set X, of size k, such that $n \in X$. In other words, we need to pick $k-1$ more elements of X, from the set $S_n \setminus \{n\}$ of size $n-1$. Therefore there are $\binom{n-1}{k-1}$ ways of doing this.

Stage 2: We have already put k elements into one part of the partition, so now we have to partition the remaining $n-k$ elements. This can be done in B_{n-k} ways.

Hence, by the multiplication principle, we have

$$|T_k| = \binom{n-1}{k-1} B_{n-k}.$$

The final result follows from the addition principle.

Example 25 *We can use this recurrence relation to compute more Bell numbers:*

$$\begin{aligned} B_4 &= \binom{3}{0}B_3 + \binom{3}{1}B_2 + \binom{3}{2}B_1 + \binom{3}{3}B_0 \\ &= 5 + 3 \times 2 + 3 + 1 = 15 \\ B_5 &= B_4 + 4B_3 + 6B_2 + 4B_1 + B_0 \\ &= 15 + 20 + 12 + 4 + 1 = 52 \end{aligned}$$

2.2 Linear recurrence relations with constant coefficients

Suppose $f_1, f_2, \ldots$ is a sequence of real numbers. Then a *recurrence relation* for the f_n is a formula which expresses f_n in terms of $f_{n-1}, f_{n-2}, \ldots, f_1$. Thus any recurrence relation enables you to calculate f_n recursively.

Some recurrence relations can be 'solved' to give a nice closed formula for f_n, but most cannot.

A k-term recurrence relation is one which expresses f_n in terms of f_{n-1}, f_{n-2}, …, f_{n-k} only. That is, it only uses the previous k terms, not all terms.

A k-term recurrence relation is *linear* if it is of the form

$$f_n = a.f_{n-1} + b.f_{n-2} + \cdots + z.f_{n-k},$$

where $a, b, \ldots, z$ may be functions of n, but do not involve any f_j. If $a, b, \ldots, z$ are constants, we say the recurrence relation has constant coefficients. The easiest recurrence relations to solve are those which are linear with constant coefficients. (Compare differential equations.)

Lecture 8, 11/10/11

Selected solutions to the first set of exercises. The first question was generally answered well, but many of you need to include more detailed explanation. Only the last part caused any real difficulty: how many sequences of five distinct numbers from $\{1,2,3,4,5,6,7\}$ contain exactly 2 of the numbers $1,2,3$? Probably the easiest way to count them is to count the *subsets* first, and then put the numbers into order. Thus there are $\binom{3}{2}=3$ ways of choosing 2 numbers from $\{1,2,3\}$, and $\binom{4}{3}=4$ ways of choosing 3 numbers from $\{4,5,6,7\}$, making 12 choices of subsets of size 5. For each of these subsets there are $5!=120$ ways of ordering it. Hence the answer is $12\times 120=1440$.

Question 2(iii) asked for the number of subsets of size 5 of $\{A,B,C,D,E,F,G,H,I\}$ which contain at least two vowels. Since there are only three vowels altogether, we get the answer by adding together the number of subsets with two vowels, and the number with three vowels. In the first case, we choose two vowels from three in $\binom{3}{2}=3$ ways, and choose three consonants from six in $\binom{6}{3}=20$ ways, making $3\times 20=60$ in all. In the second case, we take all the vowels, and choose two of the six consonants in $\binom{6}{2}=15$ ways. Hence the answer is $60+15=75$.

Question 2(iv) asked for the number of sequences of length 11 made out of the letters of MISSISSIPPI. If we first choose the place to put the M, we have $\binom{11}{1}=11$ places to put it. If we next choose two of the remaining places to put the two Ps, we have $\binom{10}{2}=45$ choices. Finally we choose 4 of the last 8 places to put the Is, in $\binom{8}{4}=70$ ways. Hence the answer is $11\times 45\times 70=34650$. Comment: if we write out the formula for these binomial coefficients, we get

$$\frac{11!}{1!10!}.\frac{10!}{2!8!}.\frac{8!}{4!4!}=\frac{11!}{1!2!4!4!}.$$

This can be interpreted by saying we have 11! orderings, but then we have counted each sequence 1!2!4!4! times, since any permutation of the two Ps, or the four Is, or the four Ss, makes no difference to the sequence.

Question 3 was deliberately hard: only one person managed to do it without a hint. We let $A=\{1,2,3,\ldots,n\}$, and $\mathcal{X}=\{B\subseteq A \mid |B| \text{ is even}\}$, and $\mathcal{Y}=\{C\subseteq A \mid |C| \text{ is odd}\}$, and we want to construct a bijection between $\mathcal{X}$ and $\mathcal{Y}$. This is easy if n is odd, and was done in lecture 4. However, it is much more difficult if n is even. The trick is to define $f:\mathcal{X}\to\mathcal{Y}$ by $f(B)=B\triangle\{1\}$. Thus $|f(B)|=|B|\pm 1$, so f really does map $\mathcal{X}$ to $\mathcal{Y}$. Moreover, f has an inverse, namely the map $g:\mathcal{Y}\to\mathcal{X}$ defined by the same formula, that is $g(C)=C\triangle\{1\}$. Hence f is a bijection, and $|\mathcal{X}|=|\mathcal{Y}|$.

Back to recurrence relations.

Example 26 • *$f_n=f_{n-1}+f_{n-2}f_{n-3}$ is a non-linear 3-term recurrence relation, because it expresses f_n in terms of the previous 3 terms f_{n-1},f_{n-2},f_{n-3};*

- *$f_n = nf_{n-1} + n^2 f_{n-2}$ is a linear 2-term recurrence relation, with non-constant coefficients;*
- *$f_n = 2f_{n-1} - f_{n-2}$ is a linear 2-term recurrence relation, with constant coefficients;*
- *$f_n = f_{n-1} + f_{n-2} + \cdots + f_1 + f_0$, should probably not be regarded as linear, since the number of terms is unbounded.*

If we are given a k-term recurrence relation for f_n, together with the first k values $f_0, f_1, \ldots, f_{k-1}$, then we can calculate successively $f_k, f_{k+1}, \ldots$, using the recurrence relation.

A general k-term linear recurrence relation is of the form

$$f_n = c_1 f_{n-1} + c_2 f_{n-2} + \cdots + c_k f_{n-k}.$$

Lemma 1 *If f_n and g_n satisfy the same k-term linear recurrence relation, then so does $h_n = Af_n + Bg_n$, for any constants A and B.*

We have

$$\begin{aligned}
f_n &= c_1 f_{n-1} + c_2 f_{n-2} + \cdots + c_k f_{n-k} \\
\Rightarrow Af_n &= c_1 Af_{n-1} + c_2 Af_{n-2} + \cdots + c_k Af_{n-k} \\
g_n &= c_1 g_{n-1} + c_2 g_{n-2} + \cdots + c_k g_{n-k} \\
\Rightarrow Bg_n &= c_1 Bg_{n-1} + c_2 Bg_{n-2} + \cdots + c_k Bg_{n-k} \\
\Rightarrow Af_n + Bg_n &= c_1(Af_{n-1} + Bg_{n-1}) + c_2(Af_{n-2} + Bg_{n-2}) + \cdots + c_k(Af_{n-k} + Bg_{n-k}) \\
\Rightarrow h_n &= c_1 h_{n-1} + c_2 h_{n-2} + \cdots + c_k h_{n-k}
\end{aligned}$$

Example 27 *If $f_n = 3f_{n-1}$, and $f_0 = 2$, write down a simple formula for f_n.*

Answer: you can probably see immediately that $f_n = 2.3^n$. In any case, it is not hard to guess that the solution should be of the form $f_n = A.x^n$ for some x and some A. Substituting in gives $A.x^n = 3A.x^{n-1}$, and cancelling $A.x^{n-1}$ gives $x = 3$. Then we find A by substituting in the initial conditions: $f_0 = A$ so $A = 2$.

Lecture 9, 13/10/11

Example 28 *If $f_n = f_{n-1} + 2f_{n-2}$ for $n \geq 2$, and $f_0 = 0$, $f_1 = 1$, find a simple formula for f_n.*

Answer: the previous example suggests we try solutions of the form $f_n = x^n$. So substitute this into the recurrence relation, to get

$$x^n = x^{n-1} + 2x^{n-2}.$$

Re-writing this as $x^n - x^{n-1} - 2x^{n-2} = 0$ gives

$$0 = x^{n-2}(x^2 - x - 2) = x^{n-2}(x+1)(x-2),$$

so $x = 0$, -1, or 2. The case $x = 0$ gives the trivial solution $f_n = 0$. The other two cases give solutions $(-1)^n$ and 2^n.

Now the lemma tells us that $A(-1)^n + B2^n$ is a solution to the recurrence relation, for any constants A and B, so let us try to put $f_n = A(-1)^n + B2^n$, and see what happens. We can substitute in the known values f_0 and f_1 to get the equations

$$\begin{aligned} 0 = f_0 &= A + B \\ 1 = f_1 &= -A + 2B \\ \Rightarrow 1 &= 3B \\ \Rightarrow B &= \frac{1}{3} \\ \Rightarrow A &= -\frac{1}{3} \end{aligned}$$

Hence the solution is

$$f_n = \frac{1}{3}(2^n - (-1)^n) = \frac{1}{3}(2^n + (-1)^{n-1}).$$

Example 29 *(Fibonacci numbers.) The Fibonacci numbers F_n are defined by the recurrence relation $F_n = F_{n-1} + F_{n-2}$ (for $n \geq 2$), and the initial conditions $F_0 = 1$, $F_1 = 1$.*

Suppose there is a solution to the recurrence relation of the form $F_n = x^n$. Substituting into the recurrence relation we get $x^n = x^{n-1} + x^{n-2}$, that is

$$x^n(x^2 - x - 1) = 0.$$

Hence $x = 0$ or $x = (1 \pm \sqrt{5})/2$. Therefore we get two independent solutions of the form $F_n = \alpha^n$ and $F_n = \beta^n$, where $\alpha = (1 + \sqrt{5})/2$ and $\beta = (1 - \sqrt{5})/2$. So by the lemma, the general solution is

$$F_n = A\alpha^n + B\beta^n,$$

for arbitrary constants A, B.

Now substitute in the initial conditions $F_0 = F_1 = 1$ to get

$$\begin{aligned} F_0 = 1 &= A + B \\ F_1 = 1 &= A\alpha + B\beta \\ \Rightarrow \alpha &= A\alpha + B\alpha \\ \Rightarrow \alpha - 1 &= B(\alpha - \beta) \\ \Rightarrow (\sqrt{5} - 1)/2 &= B\sqrt{5} \\ \Rightarrow B &= -\beta/\sqrt{5} \\ \Rightarrow A &= \alpha/\sqrt{5} \end{aligned}$$

since $\alpha - \beta = \sqrt{5}$. Hence

$$F_n = A\alpha^n + B\beta^n = \frac{1}{\sqrt{5}}(\alpha^{n+1} - \beta^{n+1}).$$

Example 30 *This method is not going to work if we have a* repeated *root of the quadratic equation, for example if* $f_n = 4f_{n-1} - 4f_{n-2}$. *So how do we solve this recurrence relation, for example with initial conditions* $f_0 = 1, f_1 = 4$*?*

In this case putting $f_n = x^n$ yields the equation $x^{n-2}(x-2)^2 = 0$, with repeated root $x = 2$, so certainly $f_n = 2^n$ is a solution to the recurrence relation. The key point is to notice that actually $f_n = n.2^n$ is also a solution: for in this case

$$\begin{aligned} f_n - 4f_{n-1} + 4f_{n-2} &= n.2^n - 4(n-1)2^{n-1} + 4(n-2)2^{n-2} \\ &= 2^n(n - 2(n-1) + (n-2)) = 0 \end{aligned}$$

So we try the general solution

$$f_n = (A + Bn)2^n.$$

Now substitute in the initial conditions, to get

$$\begin{aligned} f_0 = 1 &= A \\ f_1 = 4 &= 2A + 2B \\ \Rightarrow B &= 1 \end{aligned}$$

so the solution is $f_n = (n+1)2^n$.

General method for solving k-term linear recurrence relations with constant coefficients. Suppose we want to solve the recurrence relation

$$f_n = c_1 f_{n-1} + c_2 f_{n-2} + \cdots + c_k f_{n-k}$$

(for $n \geq k$), where $c_1, \ldots, c_k$ are constants, subject to initial values $f_0 = a_0$, $f_1 = a_1$, $\ldots$, $f_{k-1} = a_{k-1}$.

Step 1: find the roots of the characteristic equation:

$$x^k - c_1 x^{k-1} - c_2 x^{k-2} - \cdots - c_{k-1}x - c_k = 0.$$

Suppose the roots are α_1 with multiplicity m_1, and α_2 with multiplicity m_2, and so on, up to α_r with multiplicity m_r. Then $m_1 + m_2 + \cdots + m_r = k$.

Step 2: the solutions corresponding to each α_i are

$$(A_i + B_i n + C_i n^2 + \cdots + Z_i n^{m_i - 1})\alpha_i^n.$$

The number of arbitrary constants in this expression is m_i. Putting f_n equal to the sum of all of these, for $1 \leq i \leq r$, gives an expression with $m_1 + m_2 + \cdots + m_r = k$ arbitrary constants.

Step 3: substitute in the values $f_0 = a_0, \ldots, f_{k-1} = a_{k-1}$ to get k simultaneous linear equations in k unkowns $A_1, \ldots, Z_r$. Solve these to get the unique solution for f_n.

Lecture 10, 17/10/11

Example 31 *Find a formula for f_n defined by $f_n = 3f_{n-2} + 2f_{n-3}$ (for $n \geq 3$) and $f_0 = 2$, $f_1 = 0$, $f_2 = 7$.*

Answer: if $f_n = x^n$, then $x^n - 3x^{n-2} - 2x^{n-3} = 0$. Ignoring the trivial solution $x = 0$, we have

$$0 = x^3 - 3x - 2 = (x+1)^2(x-2).$$

The solution $x = 2$ gives $f_n = A2^n$, while $x = -1$ gives two independent solutions, $f_n = B(-1)^n + Cn(-1)^n$, or the general solution

$$f_n = A2^n + B(-1)^n + Cn(-1)^n.$$

Substiting $n = 0, 1, 2$ gives the three equations

$$\begin{aligned} 2 &= A + B \\ 0 &= 2A - B - C \\ 7 &= 4A + B + 2C \end{aligned}$$

which you solve in the usual way to get $A = 1$, $B = 1$, $C = 1$. Hence the solution is

$$f_n = 2^n + (n+1)(-1)^n.$$

2.3 Generating functions

Generating functions give another way of studying sequences of numbers, and often getting a simple formula for the terms of the sequence, starting from a recurrence relation, for example.

Definition 6 *If f_0, f_1, ..., is an infinite sequence of numbers, then the* generating function *$\phi(t)$ for f_n is the power series*

$$\phi(t) = \sum_{k=0}^{\infty} f_k t^k.$$

Example 32 *The Fibonacci sequence $1, 1, 2, 3, 5, 8, \ldots$ has generating function*

$$\phi(t) = 1 + t + 2t^2 + 3t^3 + 5t^4 + 8t^5 + \cdots.$$

Of course this only becomes interesting when we can get a nice formula for $\phi(t)$. In this example we can use the recurrence relation $F_n = F_{n-1} + F_{n-2}$ to get a 'functional equation' for $\phi(t)$, and solve this to get a nice formula for $\phi(t)$.

$$\phi(t) = \sum_{k=0}^{\infty} F_k t^k$$

$$\begin{aligned}
&= 1 + t + \sum_{k=2}^{\infty} F_k t^k \\
&= 1 + t + \sum_{k=2}^{\infty} (F_{k-1} + F_{k-2}) t^k \\
&= 1 + t + \sum_{k=2}^{\infty} F_{k-1} t^k + \sum_{k=2}^{\infty} F_{k-2} t^k \\
&= 1 + t + t \left(\sum_{l(=k-1)=1}^{\infty} F_l t^l \right) + t^2 \left(\sum_{m(=k-2)=0}^{\infty} F_m t^m \right) \\
&= 1 + t + t(\phi(t) - 1) + t^2 \phi(t) \\
&= 1 + t\phi(t) + t^2 \phi(t) \\
\Rightarrow (1 - t - t^2)\phi(t) &= 1 \\
\Rightarrow \phi(t) &= (1 - t - t^2)^{-1}
\end{aligned}$$

So far so good, but we can do more than this, by factorising the denominator and hence splitting this expression into partial fractions. First note that

$$1 - t - t^2 = (1 - \alpha t)(1 - \beta t),$$

where $\alpha = (1 + \sqrt{5})/2$ and $\beta = (1 - \sqrt{5})/2$. Writing

$$\phi(t) = \frac{1}{1 - t - t^2} = \frac{1}{(1 - \alpha t)(1 - \beta t)} = \frac{A}{1 - \alpha t} + \frac{B}{1 - \beta t}$$

and multiplying up by the denominator we get

$$1 = A(1 - \beta t) + B(1 - \alpha t).$$

This must hold for all values of t, and substituting $t = 1/\alpha$ gives $1 = A(\alpha - \beta)/\alpha$ so $A = \alpha/\sqrt{5}$. Similarly $B = -\beta/\sqrt{5}$. Hence

$$\phi(t) = \frac{\alpha}{\sqrt{5}} \left(\frac{1}{1 - \alpha t} \right) - \frac{\beta}{\sqrt{5}} \left(\frac{1}{1 - \beta t} \right).$$

Finally we use the Taylor expansion $(1 - x)^{-1} = 1 + x + x^2 + x^3 + \cdots$ to get

$$\phi(t) = \frac{\alpha}{\sqrt{5}} \sum_{k=0}^{\infty} \alpha^k t^k - \frac{\beta}{\sqrt{5}} \sum_{k=0}^{\infty} \beta^k t^k.$$

Comparing this with the definition of $\phi(t)$ then gives

$$F_n = \frac{\alpha^{n+1} - \beta^{n+1}}{\sqrt{5}}.$$

Power series Suppose that $\phi(t) = \sum_{k\geq 0} a_k t^k$ and $\theta(t) = \sum_{k\geq 0} b_k t^k$ are two power series. Then

1. we *define* the sum $\phi(t) + \theta(t) = \sum_{k\geq 0}(a_k + b_k)t^k$;
2. we *define* the product $\phi(t).\theta(t) = \sum_{n\geq 0} c_n t^n$ where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. This is just a formal way of saying that you multiply out the expressions as usual, and collect together the terms in each power of t:

$$(a_0+a_1t+a_2t^2+\cdots)(b_0+b_1t+b_2t^2+\cdots) = a_0b_0+(a_0b_1+a_1b_0)t+(a_0b_2+a_1b_1+a_2b_0)t^2+\cdots$$

Lecture 11, 18/10/11

3. we *define* the derivative $\phi'(t) = \sum_{k=0}^{\infty} k a_k t^{k-1} = \sum_{k=1}^{\infty} k a_k t^{k-1}$.

Example 33 *The Taylor expansion of e^t about $t = 0$ is*

$$e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{t^n}{n!}.$$

Example 34 *The Taylor expansion of $(1 + t)^r$, for fixed real $r \neq 0$, is*

$$(1 + t)^r = \sum_{n=0}^{\infty} \frac{r(r-1).\cdots.(r-n+1)}{n!} t^n.$$

Notation: write

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

even when r is not a positive integer.

Selected solutions to Exercises 2. Q.2. The wording of the question is somewhat ambiguous, but what I intended to ask was, is $n = 8$ the only solution to $\binom{n}{3} = 2\binom{n}{2}$? The answer is yes, because if

$$\frac{n(n-1)(n-2)}{3!} = 2\frac{n(n-1)}{2!}$$

then since $n \geq 3$ we can cancel the factor of $n(n-1)$ and get $n-2 = 6$, so $n = 8$.

Q.4(c). How many permutations of $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ are there consisting of three cycles of length at least 2? There are three cases, $2 + 2 + 5$, $2 + 3 + 4$, and $3 + 3 + 3$. In the case $2 + 3 + 4$, we have elements of shape

$$(x_1, x_2)(x_3, x_4, x_5)(x_6, x_7, x_8, x_9)$$

and there are 9! ways of choosing $x_1, \ldots, x_9$. But $(x_1, x_2) = (x_2, x_1)$, etc., so each such permutation has $2 \times 3 \times 4 = 24$ different names. So the number of permutations of this shape is $9!/24 = 15120$.

Now in the case $2+2+5$, this argument would give $9!/20$, but we also have to take into account the fact that $(x_1, x_2)(x_3, x_4) = (x_3, x_4)(x_1, x_2)$, which doubles the number of names of each permutation. So there are $9!/40 = 9072$ such permutations. Similarly, in the case $3+3+3$, the three 3-cycles can be permuted in any way we like, so that altogether each such permutation has $3.3.3.3! = 162$ names. Thus the number of such permutations is $9!/162 = 2240$. Adding up these three numbers gives the total

$$15120 + 9072 + 2240 = 26432.$$

Q.6(a). Calculate $\sum_{k=0}^{n} k\binom{n}{k} = \sum_{k=1}^{n} k\binom{n}{k}$.
Answer 1 (using the formula):

$$k\binom{n}{k} = \frac{k.n!}{k!(n-k)!} = \frac{n.(n-1)!}{(k-1)!(n-k)!} = n\binom{n-1}{k-1}$$

so

$$\sum_{k=1}^{n} k\binom{n}{k} = \sum_{k=1}^{n} n\binom{n-1}{k-1} = n\sum_{m=0}^{n-1}\binom{n-1}{m} = n.2^{n-1}.$$

Answer 2 (using the binomial theorem): Differentiate $(1+x)^k = \sum_{k=0}^{n}\binom{n}{k}x^k$ to get

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

and substitute $x = 1$ to get

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

Answer 3:

$$\begin{aligned}\sum_{k=0}^{n} k\binom{n}{k} &= 0.\binom{n}{0} + 1.\binom{n}{1} + 2.\binom{n}{2} + \cdots + (n-1)\binom{n}{n-1} + n\binom{n}{n} \\ &= \frac{n}{2}\binom{n}{0} + \frac{n}{2}\binom{n}{1} + \cdots + \frac{n}{2}\binom{n}{n}\end{aligned}$$

because $\binom{n}{k} = \binom{n}{n-k}$ and if we take the two corresponding terms together we have

$$k\binom{n}{k} + (n-k)\binom{n}{n-k} = n\binom{n}{k} = \frac{n}{2}\binom{n}{k} + \frac{n}{2}\binom{n}{n-k}.$$

Q.6(b). Either differentiate twice, or use the formula. In either case it is useful to note that $k^2 = k(k-1) + k$.

Lecture 12, 20/10/11

Philosophy of generating functions. In calculus, you take a nice function $f(t)$, and expand it as a power series (Taylor–Maclaurin series)

$$f(t) = f(0) + f'(0)t + \frac{f''(0)}{2!}t^2 + \cdots$$

In combinatorics, we reverse this process, and take a power series

$$f_0 + f_1 t + f_2 t^2 + \cdots$$

(where f_n is the number of 'things' of 'size' n you want to count), and try to write it as a nice function $f(t)$.

Step 1: Use a recurrence relation to get an equation for $f(t)$. For example, as we saw above, the linear recurrence relation $f_n - f_{n-1} - f_{n-2} = 0$ for Fibonacci numbers gives a linear functional equation

$$(1 - t - t^2)f(t) = \text{ something depending on the initial conditions.}$$

Step 2: Find a solution, and manipulate it into a nice form.

Step 3: Go back to the calculus to find a power series expansion for $f(t)$, which we hope will give a nice formula for the coefficients f_n.

2.4 The Catalan numbers: a case study in using generating functions

The Catalan numbers are defined by $c_1 = 1$ and the recurrence relation

$$c_n = \sum_{i=1}^{n-1} c_i c_{n-i}$$

for all $n \geq 2$. Thus we have $c_2 = c_1 c_1 = 1$, $c_3 = c_1 c_2 + c_2 c_1 = 2$, $c_4 = c_1 c_3 + c_2 c_2 + c_3 c_1 = 2 + 1 + 2 = 5$, and so on. Note that this is not a linear recurrence relation.

Now let

$$\begin{aligned} \phi(t) &= c_1 t + c_2 t^2 + c_3 t^3 + \cdots \\ &= t + t^2 + 2t^3 + 5t^4 + \cdots \\ &= \sum_{c \geq 1} c_k t^k \end{aligned}$$

and consider

$$\phi(t).\phi(t) = (c_1 t + c_2 t^2 + \cdots).(c_1 t + c_2 t^2 + \cdots)$$

$$\begin{aligned} &= c_1c_1t^2 + (c_1c_2 + c_2c_1)t^3 + (c_1c_3 + c_2c_2 + c_3c_1)t^4 + \cdots \\ &= \sum_{n\geq 2}\left(\sum_{k=1}^{n-1} c_k c_{n-k}\right)t^n \\ &= \sum_{n\geq 2} c_n t^n \qquad \text{(using the recurrence relation)} \\ &= \phi(t) - t \end{aligned}$$

So we end up with the functional equation

$$\phi(t)^2 - \phi(t) - t = 0.$$

This is just a quadratic equation in the unknown $\phi(t)$, so we can solve it to get

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

and since $\phi(0) = 0$ we deduce we must have the minus sign. Thus

$$\phi(t) = \frac{1}{2}(1 - (1-4t)^{1/2})$$

and we expand this using the Taylor series

$$(1-4t)^{1/2} = 1 + \frac{\frac{1}{2}}{1!}(-4t) + \frac{\frac{1}{2}.\frac{-1}{2}}{2!}(-4t)^2 + \cdots$$

to get

$$\begin{aligned} \phi(t) &= c_1t + c_2t^2 + c_3t^3 + \cdots \\ &= \frac{-1}{2}.\frac{\frac{1}{2}}{1!}(-4t) + \cdots + \frac{\frac{-1}{2}.\frac{1}{2}.\frac{-1}{2}.\cdots.(\frac{3}{2}-n)}{n!}(-4t)^n + \cdots \end{aligned}$$

and now we can re-write the resulting formula for c_n in various ways;

$$\begin{aligned} c_n &= \frac{-1}{2}\left(\frac{1}{2^n}\right).\frac{1.1.3.5.7.\cdots.(2n-3)}{n!}(-1)^{n-1}(-4)^n \\ &= \frac{1}{2}.2^n.\frac{1.2.3.\cdots.(2n-3)(2n-2)}{n!2.4.6.\cdots(2n-2)} \\ &= \frac{1}{2}.2^n.\frac{(2n-2)!}{n!(n-1)!2^{n-1}} \\ &= \frac{(2n-2)!}{n!(n-1)!} \\ &= \frac{1}{n}\binom{2n-2}{n-1} \end{aligned}$$

2.5 Exponential generating functions

Lecture 13, 24/11/11

Recall that the ordinary generating function $\sum_{n\geq 0} f_n t^n$ is supposed to look like a Taylor series

$$\sum_{n\geq 0} \frac{\phi^{(n)}(0)t^n}{n!}.$$

Sometimes it makes sense to look at the *exponential generating function*

$$\sum_{n\geq 0} \frac{f_n t^n}{n!}$$

instead, so that $f_n = \phi^{(n)}(0)$ rather than $f_n = \phi^{(n)}(0)/n!$.

Example 35 *If $f_n = 1$ for all n, then the ordinary generating function is*

$$\phi(t) = \sum_{n\geq 0} t^n = 1 + t + t^2 + \cdots = (1-t)^{-1},$$

but the exponential generating function is

$$\psi(t) = \sum_{n\geq 0} \frac{t^n}{n!} = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots = e^t.$$

Multiplying together ordinary generating functions:

$$\left(\sum_{k\geq 0} a_k t^k\right) \cdot \left(\sum_{j\geq 0} b_j t^j\right) = \sum_{n\geq 0} \left(\sum_{k=0}^{n} a_k b_{n-k}\right) t^n.$$

Hence, multiplying together exponential generating functions works like this:

$$\begin{aligned}
\left(\sum_{k\geq 0} \frac{a_k t^k}{k!}\right) \cdot \left(\sum_{j\geq 0} \frac{b_j t^j}{j!}\right) &= \sum_{n\geq 0} \left(\sum_{k=0}^{n} \frac{a_k}{k!} \cdot \frac{b_{n-k}}{(n-k)!}\right) t^n \\
&= \sum_{n\geq 0} \frac{1}{n!} \left(\sum_{k=0}^{n} \frac{n!}{k!(n-k)!} a_k b_{n-k}\right) t^n \\
&= \sum_{n\geq 0} \frac{d_n}{n!} t^n
\end{aligned}$$

where

$$d_n = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}.$$

Differentiating exponential generating functions. If

$$\phi(t) = \sum_{n\geq 0} \frac{a_n t^n}{n!}$$

then

$$\begin{aligned} \phi'(t) &= \sum_{n\geq 1} \frac{a_n n t^{n-1}}{n!} \\ &= \sum_{n\geq 1} \frac{a_n}{(n-1)!} t^{n-1} \\ &= \sum_{m\geq 0} \frac{a_{m+1} t^m}{m!} \end{aligned}$$

so we have just shifted the sequence (a_n) along one place.

Example: Bell numbers. Recall the definition of Bell numbers: B_n is the number of partitions of a set of size n. We saw that $B_0 = 1$, $B_1 = 1$, and for $n \geq 1$ we have the recurrence relation

$$B_n = \sum_{i=1}^{n} \binom{n-1}{i-1} B_{n-i}.$$

Now let

$$\phi(t) = \sum_{t\geq 0} \frac{B_n t^n}{n!}$$

be the exponential generating function for the Bell numbers.

Differentiating:

$$\begin{aligned} \phi'(t) &= \sum_{n\geq 1} \frac{B_n t^{n-1}}{(n-1)!} \\ &= \sum_{n\geq 1} \frac{1}{(n-1)!} \left(\sum_{i=1}^{n} \frac{(n-1)!}{(i-1)!(n-i)!} B_{n-i} \right) t^{n-1} \\ &= \sum_{m\geq 0} \left(\sum_{k=0}^{m} \frac{B_{m-k}}{k!(m-k)!} \right) t^m \end{aligned}$$

On the other hand, we calculate

$$\begin{aligned} e^t.\phi(t) &= \left(\sum_{k\geq 0} \frac{1}{k!} t^k \right) . \left(\sum_{j\geq 0} \frac{B_j}{j!} t^j \right) \\ &= \sum_{n\geq 0} \left(\sum_{k=0}^{n} \frac{B_{n-k}}{k!(n-k)!} \right) t^n \end{aligned}$$

so $\phi'(t) = e^t\phi(t)$. Now this differential equation is separable, and we obtain

$$\int \frac{\phi'(t)}{\phi(t)}dt = \int e^t dt$$

so $\ln(\phi(t)) = e^t + C$ and therefore $\phi(t) = Ae^{e^t}$. Substituting in $t = 0$ gives $\phi(0) = B_0 = 1$ so $A = e^{-1}$, and finally

$$\phi(t) = e^{e^t - 1},$$

which is a nice formula but probably not much use for actually calculating B_n.

Chapter 3

Stirling numbers

3.1 Definitions and examples

We began this course by counting *subsets* of $X = \{1, 2, 3, \ldots, n\}$. We saw that there were 2^n subsets altogether, and for each k there were $\binom{n}{k}$ subsets of size k. Therefore

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

We counted *permutations* of X, and found there were $n!$ permutations altogether. We can again divide this up according to the number of *cycles* in the permutation. Define the *unsigned Stirling number of the first kind* $u(n, k)$ to be the number of permutations of $\{1, 2, \ldots, n\}$ which have exactly k cycles (including 1-cycles). Thus

$$\sum_{k=1}^{n} u(n, k) = n!.$$

The *signed* Stirling number of the first kind is $s(n, k) = (-1)^{n-k} u(n, k)$.

We also counted *partitions* of X, and defined the Bell number B_n to be the number of partitions of X. Again, we can divide this up according to the number of parts of the partition. We define the *Stirling number of the second kind* $S(n, k)$ to be the number of partitions of $\{1, 2, \ldots, n\}$ into exactly k parts. So

$$\sum_{k=1}^{n} S(n, k) = B_n.$$

Example 36 *If $n = 3$, then $\{1, 2, 3\}$ can be partitioned into 1, 2, or 3 parts. We have $S(3, 1) = 1$, for the partition $\{\{1, 2, 3\}\}$; and $S(3, 2) = 3$, for the partitions $\{\{1, 2\}, \{3\}\}$ and $\{\{1, 3\}, \{2\}\}$ and $\{\{2, 3\}, \{1\}\}$; and $S(3, 3) = 1$, for the partition $\{\{1\}, \{2\}, \{3\}\}$.*

More generally, we see that $S(n, 1) = 1$ and $S(n, n) = 1$ for all n.

Example 37 *Let us calculate* $u(7,3)$*, the number of permutations of* $\{1,2,3,4,5,6,7\}$ *which are the product of three disjoint cycles. Now the lengths of these cycles add up to* 7*, which can happen as* $5+1+1$ *or* $4+2+1$ *or* $3+3+1$ *or* $3+2+2$*. In the first case there are* $7!/5.2! = 504$ *such permutations; in the second there are* $7!/4.2 = 630$*; in the third there are* $7!/3.3.2! = 280$*; and in the last there are* $7!/3.2.2.2! = 210$*. This makes* $504+630+280+210 = 1624$ *altogether. Thus* $u(n,k) = 1684$*, and because* $n-k = 4$ *is even,* $s(n,k) = 1684$*.*

Example 38 *We already know that the number of permutations consisting of* n *cycles (necessarily of length* 1*) is just* 1*, so* $s(n,n) = u(n,n) = 1$*. At the other extreme, the number of permutations consisting of a single cycle is* $n!/n = (n-1)!$*. So* $u(n,1) = (n-1)!$ *and* $s(n,1) = (-1)^{n-1}(n-1)!$*.*

3.2 Recurrence relations for Stirling numbers

First let us look for a recurrence relation for the Stirling numbers of the second kind, $S(n,k)$, analogous to the relation

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

for the binomial coefficients. In other words, we want to reduce n, but we don't care what happens to k.

So, in going from partitions of $\{1,2,\ldots,n-1\}$ to partitions of $\{1,2,\ldots,n\}$, we distinguish two cases, according to what happens to the new element n: either

1. $\{n\}$ is one of the parts of the partition; or

2. n is confined in a part with something else.

In the first case, we have to partition the remaining $n-1$ elements into $k-1$ parts, so there are $S(n-1,k-1)$ partitions of this form.

In the second case, ignoring the element n for the moment, we have to partition $\{1,2,\ldots,n-1\}$ into k parts, which can be done in $S(n-1,k)$ ways. Then we have to put the element n into one of the k parts, which can be done in k ways. Thus there are $kS(n-1,k)$ partitions of this type.

Hence, by adding these two cases together, we obtain the recurrence relation

$$S(n,k) = S(n-1,k-1) + kS(n-1,k).$$

A similar argument gives a recurrence relation for the Stirling numbers of the first kind. This time we want to count permutations of $\{1,2,\ldots,n\}$ which consist of k cycles. Again we divide into two cases, according to whether (n) is a cycle on its own, or whether n is in a cycle with some other elements. In the first case,

we just need to count the permutations of $\{1, 2, \ldots, n-1\}$ which consist of $k-1$ cycles. The number of these is $u(n-1, k-1)$.

In the second case, we begin by counting the number of permutations of $\{1, 2, \ldots, n\}$ which consist of k cycles, and then consider how to insert n into the permutation. The first step can be accomplished in $u(n-1, k)$ ways, and the second in $n-1$ ways (compare the argument for Bell numbers). Hence the number of possibilities in this case is $(n-1)u(n-1, k)$. Therefore the unsigned Stirling numbers of the first kind satisfy the recurrence relation

$$u(n, k) = u(n-1, k-1) + (n-1)u(n-1, k).$$

Putting in the signs gives the recurrence relation

$$s(n, k) = s(n-1, k-1) - (n-1)s(n-1, k)$$

for the signed Stirling numbers of the first kind.

3.3 Two-variable generating functions

The binomial theorem can be regarded as a theorem about a generating function for the binomial coeffiecients. If we fix n and let k vary, then the generating function for the binomial coefficients is

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

On the other hand, if we fix k and let n vary, the generating function is

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

So why don't we consider a generating function allowing both parameters to vary? Thus:

$$\begin{aligned} \sum_{n\geq 0}\sum_{k\geq 0} \binom{n}{k} x^n y^k &= \sum_{n\geq 0} x^n (1+y)^n \\ &= (1 - x(1+y)^{-1}. \end{aligned}$$

Now let us explore the same ideas for Stirling numbers of the first kind. Can we find a simple formula for

$$\sum_{k=1}^{n} s(n, k) x^k?$$

Example 39 *For $n = 3$, we calculate $s(3,1) = 2$, $s(3,2) = -3$, and $s(3,3) = 1$, so that*

$$\sum_{k=1}^{3} s(3,k)x^k = 2x - 3x^2 + x^3 = x(x-1)(x-2).$$

This suggests the following:

Theorem 8

$$\sum_{k=1}^{n} s(n,k)x^k = x(x-1)(x-2)\cdots(x-k+1).$$

Notation: we write $(x)_k$ for the expression $x(x-1)\cdots(x-k+1)$, the so-called *falling factorial* which is a product of k factors.

Proof of the theorem was not given in the lectures. The corresponding result for $S(n,k)$ looks a bit different:

Theorem 9

$$\sum_{i=1}^{n} S(n,k)(x)_k = x^n.$$

Proof: by induction on n. The case $n = 1$ is just $s(1,1)x = x$, which is true because $s(1,1) = 1$. So assume it is true for $n-1$, that is

$$\sum_{k=1}^{n-1} S(n-1,k)(x)_k = x^{n-1}.$$

Now we can use the recurrence relation as follows:

$$\begin{aligned}
x^n &= x.x^{n-1} \\
&= x\sum_{k=1}^{n-1} S(n-1,k)(x)_k \\
&= \sum_{k=1}^{n-1} S(n-1,k)(x)_k(x-k+k) \\
&= \sum_{k=1}^{n-1} S(n-1,k)(x)_{k+1} + \sum_{k=1}^{n-1} kS(n-1,k)(x)_k \\
&= \sum_{m=2}^{n} S(n-1,m-1)(x)_m + \sum_{k=1}^{n} kS(n-1,k)(x)_k \\
&= \sum_{k=2}^{n} S(n-1,k-1)(x)_k + \sum_{k=1}^{n} kS(n-1,k)(x)_k \\
&= \sum_{k=1}^{n} (S(n-1,k-1) + kS(n-1,k))(x)_k
\end{aligned}$$

$$= \sum_{k=1}^{n} S(n,k)(x)_k$$

[Take care to check that the terms we add in when we change the range of summation are all 0, since $S(n-1,0) = S(n-1,n) = 0$.]

The last two theorems can be interpreted by saying that the two matrices of Stirling numbers are the transition matrices between two bases of the vector space of polynomials divisibe by x, namely the 'natural' basis x, x^2, x^3 ..., and the basis x, $x(x-1)$, $x(x-1)(x-2)$, ... Hence these matrices are inverses of each other.

Chapter 4

The principle of inclusion and exclusion

4.1 P.I.E.

Example 40 *Let A_1, A_2 be subsets of a set X. Then*

$$|A_1 \cup A_2| = |A_1| + |A_2| - |A_1 \cap A_2|$$

so

$$|X \setminus (A_1 \cup A_2)| = |X| - |A_1| - |A_2| + |A_1 \cap A_2|.$$

The principle of inclusion/exclusion (P.I.E.) is a generalisation of these equations to n subsets.

Theorem 10 *Let $A_1, A_2, \ldots, A_n$ be subsets of X. Then*

$$\begin{aligned}
\left|\bigcup_{i=1}^{n} A_i\right| &= |A_1 \cup A_2 \cup \cdots \cup A_n| \\
&= |A_1| + |A_2| + \cdots + |A_n| \\
&\quad -(|A_1 \cap A_2| + |A_1 \cap A_3| + \cdots + |A_{n-1} \cap A_n|) \\
&\quad +(|A_1 \cap A_2 \cap A_3| + \cdots + |A_{n-2} \cap A_{n-1} \cap A_n|) \\
&\quad -\cdots + (-1)^{n-1}|A_1 \cap A_2 \cap \cdots \cap A_n| \\
&= \sum_{\emptyset \neq I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|-1} \left|\bigcap_{i \in I} A_i\right|
\end{aligned}$$

Example 41 *The case $n = 3$ is*

$$|A_1 \cup A_2 \cup A_3| = |A_1| + |A_2| + |A_3| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3|.$$

Proof. The key to proving this is to realise that on the left-hand side we are counting each element of $|A_1 \cup A_2 \cup \cdots \cup A_n|$ exactly once, while on the right-hand side we are counting each element lots of times, with signs attached. We need to show that on the right-hand side the minus signs almost balance out the plus signs, but with one plus sign left over.

So pick any $x \in \bigcup_{i=1}^{n} A_i$, and consider how many times it is counted. First observe that x is in some of the A_i, but not necessarily all of them, so let J be the corresponding set of subscripts, that is

$$J = \{j \mid x \in A_j\}.$$

Now if $I \subseteq J$, then $x \in \bigcap_{i \in J} A_i \subseteq \bigcap_{i \in I} A_i$. (Think about this: the first intersection is an intersection of *more* sets than the second, so is a *smaller* set.) More formally, since $B \cap C \subseteq B$ for any sets B, C we have:

$$x \in \bigcap_{i \in J} A_i = (\bigcap_{i \in I} A_i) \cap (\bigcap_{i \in J \setminus I} A_i) \subseteq \bigcap_{i \in I} A_i.$$

On the other hand, if $I \not\subseteq J$, then there is some $i \in I \setminus J$, and then $x \notin A_i$, so $x \notin \bigcap_{i \in I} A_i$. Putting these two cases together, we see that $x \in \bigcap_{i \in I} A_i$ if and only if $I \subseteq J$.

So the terms on the right-hand side in which x is counted are precisely the ones where $I \subseteq J$. For each such subset I, the element x is counted either $+1$ or -1 times, according to the formula. So the total number of times x is counted is

$$\sum_{\emptyset \neq I \subseteq J} (-1)^{|I|-1}.$$

Where have we seen something like this before? We are counting $+1$ for the subsets I of *odd* size, and we are counting -1 for the subsets I of *even* size (apart from the empty set). So what is the answer? (Hint: binomial theorem.)

Well, we know that the total number of subsets (of J) of even size is equal to the total number of subsets of odd size. So if we had the empty set in the sum, it would cancel out and give 0. But we don't, so we have to subtract off the contribution which the empty set would make (which is -1), giving the answer $+1$, as required.

Notation. It is sometimes helpful to use the notation $A_I = \bigcap_{i \in I} A_i$ to avoid complicated expressions. But only do this if you are absolutely clear what it means! For example, if you can't remember whether it is $\cap$ or $\cup$, then you should not be using it. It is useful also to extend the notation to $A_\emptyset = X$, the relevant universal set.

Corollary 3 *With the above notation, for subsets A_i of a set X, we have*

$$\left| X \setminus \bigcup_{i=1}^{n} A_i \right| = \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} |A_I|.$$

Proof.

$$\begin{aligned}
\left|X - \bigcup_{i=1}^{n} A_i\right| &= |X| - \left|\bigcup_{i=1}^{n} A_i\right| \\
&= |X| - \sum_{\emptyset \neq I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|-1} \left|\bigcap_{i \in I} A_i\right| \\
&= |A_\emptyset| + \sum_{\emptyset \neq I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} |A_I| \\
&= \sum_{I \subseteq \{1,2,\ldots,n\}} (-1)^{|I|} |A_I|
\end{aligned}$$

Example 42 *How many primes are there between* 1 *and* 100*?*

Answer: suppose $x \in X = \{1, 2, \ldots, 100\}$ is *not* prime. Then $x = yz$ where y is prime and $y < z$. Hence $y < \sqrt{100} = 10$, so $y = 2, 3, 5$ or 7.

Now let $A_i = \{x \mid 1 \leq x \leq 100 \text{ and } i \text{ divides } x\}$, for $i = 2, 3, 5, 7$. Then the set of all primes in X is

$$\left(X \setminus \{1\} \setminus \bigcup_{i \in \{2,3,5,7\}} A_i\right) \cup \{2, 3, 5, 7\}.$$

Now we compute the sizes of all the intersections of the A_i:

$$\begin{aligned}
|A_2| &= \lfloor \frac{100}{2} \rfloor = 50 \\
|A_3| &= \lfloor \frac{100}{3} \rfloor = 33 \\
|A_5| &= \lfloor \frac{100}{5} \rfloor = 20 \\
|A_7| &= \lfloor \frac{100}{7} \rfloor = 14 \\
|A_2 \cap A_3| &= \lfloor \frac{100}{6} \rfloor = 16 \\
|A_2 \cap A_5| &= \lfloor \frac{100}{10} \rfloor = 10 \\
|A_2 \cap A_7| &= \lfloor \frac{100}{14} \rfloor = 7 \\
|A_3 \cap A_5| &= \lfloor \frac{100}{15} \rfloor = 6 \\
|A_3 \cap A_7| &= \lfloor \frac{100}{21} \rfloor = 4 \\
|A_5 \cap A_7| &= \lfloor \frac{100}{35} \rfloor = 2 \\
|A_2 \cap A_3 \cap A_5| &= \lfloor \frac{100}{30} \rfloor = 3
\end{aligned}$$

$$\begin{aligned} |A_2 \cap A_3 \cap A_7| &= \lfloor \frac{100}{42} \rfloor = 2 \\ |A_2 \cap A_5 \cap A_7| &= \lfloor \frac{100}{70} \rfloor = 1 \\ |A_3 \cap A_5 \cap A_7| &= \lfloor \frac{100}{105} \rfloor = 0 \\ |A_2 \cap A_3 \cap A_5 \cap A_7| &= \lfloor \frac{100}{210} \rfloor = 0 \end{aligned}$$

So by P.I.E. we have

$$\begin{aligned} |A_2 \cup A_3 \cup A_5 \cup A_7| &= (50 + 33 + 20 + 14) - (16 + 10 + 7 + 6 + 4 + 2) + (3 + 2 + 1) \\ &= 117 - 45 + 6 = 78. \end{aligned}$$

Hence the number of primes in X is $100 - 78 - 1 + 4 = 25$.

The following corollary to PIE is often useful if all the sets A_i 'look the same'.

Corollary 4 *Suppose that $A_1, A_2, \ldots, A_n$ are subsets of X, and assume that for every k with $1 \leq k \leq n$ and for every $I \subseteq \{1, 2, \ldots, n\}$ with $|I| = k$ we have*

$$\left| \bigcap_{i \in I} A_i \right| = a_k.$$

Then the number of elements in none of the A_i is

$$\left| X \setminus \bigcup_{i \in I} A_i \right| = \sum_{k=0}^{n} (-1)^k \binom{n}{k} a_k.$$

Proof: If $|I| = k$, then the contribution from I to the sum in PIE is $(-1)^k a_k$. Adding this up over all the sets of size k gives $\binom{n}{k}(-1)^k a_k$. Finally adding up over all k gives the result.

4.2 Counting surjections

Suppose S and T are sets with $|S| = n$ and $|T| = k$. Recall that the number of functions from S to T is k^n (by the multiplication principle). We also computed the number of injections from S to T to be $k(k-1)\cdots(k-n+1)$, which is 0 if $n > k$, and the number of bijections, which is $n!$ if $n = k$ and 0 otherwise. But we did not count the number of surjections.

We will now look at two ways to count surjections, one directly using PIE, and one using partitions (and Stirling numbers of the second kind). First consider an example.

Example 43 *If $|S| = n = 4$ and $T = \{1, 2, 3\}$, so that $|T| = k = 3$, how many surjections are there from S to T?*

First solution: the total number of functions is $3^4 = 81$. Let A_1 be the set of functions which do not take the value 1: then $|A_1| = 2^4 = 16$. Similarly for the sets A_2 and A_3 of functions which do not take the value 2 (respectively 3). Now $A_1 \cap A_2$ is the set of functions which take neither the value 1 nor the value 2: there is only one such function. Similarly, $|A_i \cap A_j| = 1$ in all cases. Finally, $A_1 \cap A_2 \cap A_3 = \emptyset$, as every function must take *some* value. Therefore by the corollary to PIE, the number of surjections is

$$81 - 3 \times 16 + 3 \times 1 = 36.$$

Second solution: We first partition S into three parts, in $S(4,3) = 6$ ways. Then, for each part of the partition, we decide which of the three elements of T to map the elements of this part to. Since each part of S has to map to a different element of T, we have 3! ways of doing this, making $6 \times 3! = 36$ ways altogether.

Now let's do the general case. First we use P.I.E., in effect counting the number of functions which are *not* surjections. To set up the notation, let $T = \{t_1, t_2, \ldots, t_k\}$, and let X be the set of *all* functions from S to T, so that $|X| = k^n$. For each i in the range $1 \leq i \leq k$, let A_i be the set of functions which do not take the value t_i, that is

$$A_i = \{f \in X \mid \text{for all } s \in S, f(s) \neq t_i\}.$$

This means that the set of surjective maps is just $X \setminus \bigcup_{i=1}^{k} A_i$.

In order to use P.I.E., we need to compute the sizes of the intersections of any number of the A_i. First of all, what is $|A_i|$? Well, it is simply the number of functions which map to $T \setminus \{t_i\}$, so it is $(k-1)^n$. Now what is $|A_i \cap A_j|$, if $i \neq j$? Again, it is the number of functions which map to $T \setminus \{t_i, t_j\}$, so it is $(k-2)^n$.

More generally, the intersection of m of the sets A_i is the set of functions which do not map onto any of the corresponding m elements of T, and the number of such functions is $(k-m)^n$. Now if we add this up over *all* ways of choosing exactly m of the sets to intersect, then we get

$$\binom{k}{m}(k-m)^n.$$

Hence, substituting into P.I.E. we get that the total number of surjections is

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

Stirling numbers of the second kind. There is a close connection between surjections and partitions. Suppose that $|S| = n$ and $|T| = k$ as above, and say $S = \{s_1, s_2, \ldots, s_n\}$ and $T = \{t_1, t_2, \ldots, t_k\}$. Given a surjection $f : S \to T$, we can construct a partition $\{S_1, S_2, \ldots, S_k\}$ of S by defining

$$S_i = \{s \in S \mid f(s) = t_i\}.$$

Is it obvious that this is a partition? It should be, but if not, let's check it:

- the fact that f is a surjection implies that $S_i \neq \emptyset$;
- if $i \neq j$ then $t_i \neq t_j$ and there is no s with $f(s) = t_i$ and $f(s) = t_j$, so $S_i \cap S_j = \emptyset$;
- every element $s \in S$ is mapped to *some* t_i, so the union of all the S_i is the whole of S.

But it is not just a partition, it is an *ordered* partition, because the parts $S_1, \ldots, S_k$ come in a particular order, corresponding to our chosen order of the elements of T.

Conversely, given an ordered partition $\{S_1, S_2, \ldots, S_k\}$ of S, we can define the corresponding surjection $f : S \to T$ by $f(s) = t_i$ if and only if $s \in S_i$.

In other words we have shown that there is a one-to-one correspondence between the set of *surjections* $f : S \to T$ on the one hand, and the set of *ordered partitions of S into k parts* on the other.

Now the number of *unordered* partitions of S into k parts is defined to be the Stirling number $S(n, k)$, so the number of *ordered* partitions of S into k parts is $k!S(n, k)$.

So we have two different formulae for the number of ordered partitions, so they must be equal:

$$k!S(n,k) = \sum_{m=0}^{k} (-1)^m \binom{k}{m} (k-m)^n$$

and therefore

$$S(n,k) = \frac{1}{k!} \sum_{m=0}^{k} (-1)^m \binom{k}{m} (k-m)^n.$$

4.3 Derangements

A *derangement* is a permutation which has no fixed point (that is, no cycles of length 1). More formally, a derangement of a set S is a bijection $f : S \to S$ with the property that for all $s \in S$, $f(s) \neq s$.

Example 44 *If* $S = \{1, 2, 3, 4, 5\}$ *and* $f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 4 & 3 & 5 \end{pmatrix} = (1, 2)(3, 4)(5)$ *then f is not a derangement, because $f(5) = 5$.*

But $g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix} = (1, 2, 3)(4, 5)$ *is a derangement.*

Example 45 *Suppose you have a set of n letters, and n corresponding envelopes, addressed to different people. If you put the letters into envelopes at random, what is the probability that nobody gets the correct letter?*

Answer: it is the proportion of permutations of n elements which are derangements. We now investigate what this proportion is.

We can count derangements in a similar way to the way we counted surjections, using P.I.E.; that is, we count instead the permutations which *do* have fixed points. So let's begin by setting up the notation. Let $T = \{t_1, t_2, \ldots, t_n\}$, and let X be the set of all permutations of T. Formally:

$$X = \{f : T \to T \mid f \text{ is a bijection}\}.$$

For each i in the range $1 \leq i \leq n$, let A_i be the set of all permutations which fix i. Formally:

$$A_i = \{f \in X \mid f(i) = i\}.$$

Now we see that the set of all permutations which fix at least one point is exactly the union of all the A_i. So if we are going to use PIE then we need to count the elements in the various intersections of some of the A_i.

First, what is $|A_i|$? Well, A_i consists of all the permutations f which fix i. That means $f(i) = i$, but there is no restriction on $f(x)$ for any of the other $n-1$ values of x. So $|A_i| = (n-1)!$. Similarly, if $i \neq j$, then $|A_i \cap A_j|$ is the number of permutations which fix i and j, and permute the other $n-2$ points in any way at all: therefore $|A_i \cap A_j| = (n-2)!$.

More generally, the intersection of any m of the sets A_i has size $(n-m)!$, because it consists of those permutations which fix a specified set of m elements, and permute the rest. Since there are $\binom{n}{m}$ ways of choosing the m elements to fix, we get a total contribution of $(-1)^{m-1}\binom{n}{m}(n-m)!$ to the formula for $|\bigcup A_i|$. We have

$$|\bigcup_{i=1}^{n} A_i| = \sum_{m=1}^{n} (-1)^{m-1} \binom{n}{m} (n-m)!$$

and therefore

$$\begin{aligned} |X \setminus \bigcup_{i=1}^{n} A_i| &= \sum_{m=0}^{n} (-1)^m \binom{n}{m} (n-m)! \\ &= \sum_{m=0}^{n} (-1)^m \frac{n!}{m!} \\ &= n! \sum_{m=0}^{n} \frac{(-1)^m}{m!} \end{aligned}$$

Now does this series remind you of anything? What would we get if we took the sum to infinity? It is the power series expansion of e^{-1}. So the number of

derangements of a set of size n is approximately $n!/e$. (So the *proportion* of all permutations which are derangements is approximately $1/e$.)

How good is this approximation? What is the error? It is

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

Now the terms on the right-hand side are smaller than $1/n^{m-n}$ in absolute value, and decreasing in absolute value, and alternate in sign, so the absolute value of their sum is less than the first term, which is $1/n$. In other words the error is remarkably small, and the number of derangements of a set of size n is the *nearest integer* to $n!/e$, as long as $n \geq 2$. (Indeed, you can easily check that it also works for $n = 1$, where there are no derangements.)

A recurrence relation. A second approach to counting derangements is to use a recurrence relation. Let d_n denote the number of derangements of a set of size n. Suppose that π is a derangement of $\{1, 2, \ldots, n\}$. Then π maps some element i ($i \neq n$) to n, that is $\pi(i) = n$. Note that there are $n-1$ choices for i. Now we divide into two cases depending on whether $\pi(n) = i$ or not.

Case 1: $\pi(n) = i$. In this case, π consists of the 2-cycle (i, n), together with a *derangement* π^* of the remaining $n-2$ points. Hence there are d_{n-2} ways to choose π^*. So altogether there are $(n-1)d_{n-2}$ derangements π in this case.

Case 2: $\pi(n) = j \neq i$. So in cycle notation $\pi = (\ldots, i, n, j)(\ldots)\ldots$. Now if we just remove n from this we get a derangement $\pi' = (\ldots, i, j)(\ldots)\ldots$ of $\{1, 2, \ldots, n-1\}$. There are d_{n-1} choices for π'. Now to get back from π' to π we need to put back n. There are $n-1$ places we can put it (we can put it after any of the elements $1, 2, 3, \ldots, n-1$). Hence there are $(n-1)d_{n-1}$ derangements π in this case.

So altogether, $d_n = (n-1)(d_{n-1} + d_{n-2})$. In fact, this 2-term recurrence relation can be simplified to the 1-term recurrence relation

$$d_n = nd_{n-1} + (-1)^n.$$

Proof: by induction on n. To check the case $n = 2$ we verify that $d_2 = 1$, and $d_1 = 0$, so $nd_{n-1} + (-1)^n = 2 \times 0 + (-1)^2 = 1$.

Now suppose $d_{n-1} = (n-1)d_{n-2} + (-1)^{n-1}$, so that

$$\begin{aligned} d_n &= (n-1)d_{n-1} + (n-1)d_{n-2} \\ &= (n-1)d_{n-1} + d_{n-1} - (-1)^{n-1} \\ &= nd_{n-1} + (-1)^n \end{aligned}$$

as required.

Dividing this through by $n!$ gives us

$$\frac{d_n}{n!} = \frac{nd_{n-1}}{n!} + \frac{(-1)^n}{n!}$$

$$\begin{aligned} &= \frac{d_{n-1}}{(n-1)!} + \frac{(-1)^n}{n!} \\ &= \sum_{r=0}^{n} \frac{(-1)^r}{r!} \end{aligned}$$

by induction on n again. So we again obtain the formula

$$d_n = n! \sum_{r=0}^{n} \frac{(-1)^r}{r!}.$$

The exponential generating function. A third approach to counting derangements is to use a generating function. So let d_n be the number of derangements of a set of n elements, and let the exponential generating function be

$$D(x) = \sum_{n \geq 0} \frac{d_n x^n}{n!}.$$

Now if k is a permutation with exactly k fixed points, then it is a derangement of the remaining $n-k$ points. Hence the total number of permutations with exactly k fixed points is $\binom{n}{k} d_{n-k}$. Summing over all possible values of k, we have $n! = \sum_{k=0}^{n} \binom{n}{k} d_{n-k}$, so

$$1 = \sum_{k=0}^{n} \frac{1}{k!} \frac{d_{n-k}}{(n-k)!}.$$

Putting the two sides of this equation as the coefficients in an ordinary generating function we have

$$\begin{aligned} (1-x)^{-1} &= \sum_{n \geq 0} x^n \\ &= \sum_{n \geq 0} \left(\sum_{k=0}^{n} \frac{1}{k!} \frac{d_{n-k}}{(n-k)!} \right) x^n \\ &= \left(\sum_{k \geq 0} \frac{x^k}{k!} \right) \cdot \left(\sum_{r \geq 0} \frac{d_r x^r}{r!} \right) \\ &= e^x D(x) \end{aligned}$$

Hence $D(x) = e^{-x}(1-x)^{-1}$. Multiplying out the power series expansions of e^{-x} and $(1-x)^{-1}$ gives us the same formula for d_n as we obtained by the other two methods.

Chapter 5

Systems of distinct representatives

Definition 7 *If $A_1, A_2, \ldots, A_n$ are sets, then a* system of distinct representatives *(SDR) for the A_i is a set of ordered pairs*

$$f = \{(a_1, A_1), (a_2, A_2), \ldots, (a_n, A_n)\}$$

with $a_i \in A_i$ for each i, and $a_i \neq a_j$ whenever $i \neq j$.

Thus a_i is a *representative* of the set A_i, and different sets have different representatives.

Note: we do not insist that the A_i are distinct. This is why we do not talk about the *set* $\{A_1, A_2, \ldots, A_n\}$.

Example 46 *Suppose the Students' Union has a number of Clubs, and wants to choose a Committee consisting of one member of each Club, but insists that no individual can represent more than one Club.*

Example 47 $A_1 = \{1, 2, 4\}$, $A_2 = \{1, 3, 4\}$, $A_3 = \{1, 2\}$, $A_4 = \{1, 2\}$, $A_5 = \{1, 5\}$. *Then*

$$\{(4, A_1), (3, A_2), (2, A_3), (1, A_4), (5, A_5)\}$$

is a SDR for these sets.

Example 48 $B_1 = \{1, 2\}$, $B_2 = \{3, 4, 5\}$, $B_3 = \{2, 3\}$, $B_4 = \{1, 3\}$, $B_5 = \{1, 2, 3\}$. *Do these sets have a SDR?*

Answer: we see that $B_1 \cup B_3 \cup B_4 \cup B_5 = \{1, 2, 3\}$. Now we need four different representatives for these four sets, but we only have three elements to choose from. So there is no SDR.

We conclude from this argument that if $A_1, \ldots, A_n$ has a SDR, then for any choice of k of the A_i, the union of these k sets must have at least k elements. What is rather remarkable is that this condition is not only necessary, it is also sufficient.

Lemma 2 *If $A_1, A_2, \ldots, A_n$ has a SDR, then for every $I \subseteq \{1, 2, \ldots, n\}$ we have*

$$\left| \bigcup_{i \in I} A_i \right| \geq |I|.$$

Proof: This is more or less obvious from the preceding discussion. Suppose we have a SDR

$$\{(a_1, A_1), \ldots, (a_n, A_n)\},$$

and suppose that $I \subseteq \{1, 2, \ldots, n\}$. Then for each $i \in I$ we have $a_i \in A_i \subseteq \bigcup_{i \in I} A_i$, so $\{a_i \mid i \in I\} \subseteq \bigcup_{i \in I} A_i$. Now we know that $a_i \neq a_j$ whenever $i \neq j$, so $|\{a_i \mid i \in I\}| = |I|$. Hence

$$\left| \bigcup_{i \in I} A_i \right| \geq |\{a_i \mid i \in I\}| = |I|.$$

Definition 8 *If the sets $A_1, \ldots, A_n$ have the property that $|\bigcup_{i \in I} A_i| \geq |I|$ for all $I \subseteq \{1, 2, \ldots, n\}$, we say that they* satisfy Hall's condition.

So we have seen that Hall's condition is a necessary condition for the existence of an SDR. We now show it is also a sufficient condition.

Theorem 11 *(Hall's Marriage Theorem) Let $A_1, A_2, \ldots, A_n$ be sets. Then they have an SDR if and only if they satisfy Hall's condition.*

Proof: The 'only if' direction was proved in the Lemma.

Now suppose that Hall's condition is satisfied. We construct a SDR by induction on n. The case $n = 1$ is easy: Hall's condition says $|A_1| \geq 1$, so we can choose any $a_1 \in A_1$ as a representative. Then $\{(a_1, A_1)\}$ is a SDR.

So we may assume $n \geq 2$, and that the result is true for any collection of at most $n - 1$ sets. Now the proof splits into two cases depending on whether there is always room to spare, or whether there is some collection of sets with only just enough elements to go round. That is, in case 1, we always have $|\bigcup_{i \in I} A_i| > |I|$ for all I except $\emptyset$ and $\{1, 2, \ldots, n\}$. And in case 2, we have some *critical set* J with $|\bigcup_{j \in J} A_j| = |J|$.

Case 1: no subset is critical. In this case, we choose any $a_n \in A_n$, and remove a_n from all the other sets. That is, define $A'_i = A_i \setminus \{a_n\}$. Then we easily see that $A'_1, A'_2, \ldots, A'_{n-1}$ still satisfies Hall's condition (because we have removed at most one element from each of the unions under consideration). So, by induction, we can choose a SDR for the A'_i, say

$$\{(a_1, A'_1), \ldots, (a_{n-1}, A'_{n-1})\}$$

and then

$$\{(a_1, A_1), \ldots, (a_{n-1}, A_{n-1}), (a_n, A_n)\}$$

is a SDR for the A_i.

Case 2: there is a critical subset J. Then we first choose a SDR for the sets A_j with $j \in J$. (By induction, we can do this, as $|J| < n$.) Suppose this SDR is $\{(a_j, A_j) \mid j \in J\}$. Then because J is a critical set we have

$$\bigcup_{j \in J} A_j = \{a_j \mid j \in J\}$$

and this set has size $|J|$.

We now remove these $|J|$ elements from all the remaining sets, and use induction again. That is, we let $K = \{1, 2, \ldots, n\} \setminus J$, and for each $k \in K$ let

$$A_k^* = A_k \setminus (\bigcup_{j \in J} A_j).$$

Choose any $I \subseteq K$. Then

$$\begin{aligned} \left|\bigcup_{i \in I} A_i^*\right| &= \left|\bigcup_{i \in I}(A_i^* \cup (\bigcup_{j \in J} A_j))\right| - \left|\bigcup_{j \in J} A_j\right| \\ &= \left|\bigcup_{i \in I \cup J} A_i\right| - \left|\bigcup_{j \in J} A_j\right| \\ &\geq |I \cup J| - |J| \\ &= |I| + |J| - |J| = |I| \end{aligned}$$

so the A_k^* satisfy Hall's condition, so have an SDR, say

$$\{(b_k, A_k^*) \mid k \in K\}.$$

Then

$$\{(a_j, A_j), (b_k, A_k) \mid j \in J, k \in K\}$$

is a SDR for the A_i.

Here are a couple of examples to illustrate the two cases in the proof.

Lecture 20, 22/11/11

Example 49 *Let* $A_1 = \{1, 2, 3\}$, $A_2 = \{1, 3, 4\}$, $A_3 = \{2, 3, 4\}$, $A_4 = \{1, 2\}$. *Now it is easy to see that*

(a) each A_i *has* $|A_i| > 1$*;*

(b) each union $A_i \cup A_j$ *has* $|A_i \cup A_j| \geq 3 > 2$*;*

(c) $|A_i \cup A_j \cup A_k| \geq 4 > 3$*;*

Hence we are in case 1, and we can choose any $a_4 \in A_4$*, say* $a_4 = 1$*. Then we remove the element* 1 *from all the other sets to get* $A_1' = \{2, 3\}$*,* $A_2' = \{3, 4\}$*,* $A_3' = \{2, 3, 4\}$*. Now because we have only removed one element from the sets, the sizes of the various unions go down by at most* 1*, and Hall's condition is still satisfied. So we can find a SDR of the* A_i'*, say*

$$\{(2, A_1'), (3, A_2'), (4, A_3')\}.$$

Then

$$\{(2, A_1), (3, A_2), (4, A_3), (1, A_4)\}$$

is a SDR for the A_i*.*

Example 50 *Let* $B_1 = \{1, 2, 3\}$*,* $B_2 = \{2, 3, 4\}$*,* $B_3 = \{2, 4\}$*,* $B_4 = \{1, 4, 5\}$*,* $B_5 = \{3, 4\}$*. Now we notice that* $B_2 \cup B_3 \cup B_5 = \{2, 3, 4\}$*, so that* $\{2, 3, 4\}$ *is a critical set. By induction we find a SDR for the sets* B_2, B_3, B_5*, say*

$$\{(2, B_2), (4, B_3), (3, B_5)\}.$$

Then we remove the elements $2, 3, 4$ *from the remaning sets to get* $B_1^* = \{1\}$*,* $B_4^* = \{1, 5\}$*, and pick a SDR for these sets, which can only be*

$$\{(1, B_1^*), (5, B_4^*)\}.$$

Then putting these SDRs together we get an SDR for the B_i*, that is*

$$\{(2, B_2), (4, B_3), (4, B_5), (1, B_1), (5, B_4)\}.$$

Remark: The idea of a 'critical set', used in this proof, may be familiar to you from sudoku. Specifically, you can often find a set of k boxes in a row, or column, or block, with the property that there are only k numbers available for all these boxes together. Then you know that the remaining numbers must be distributed amongst the remaining boxes.

Remark: In practice, checking Hall's condition can be very time-consuming, because you need to check all $2^n - 2$ non-trivial subsets of $\{1, 2, \ldots, n\}$. So it is really only useful when you can prove that Hall's condition is satisfied, for some theoretical reason.

Now there are a number of useful corollaries of Hall's marriage theorem (some of which we may use in later chapters). For example:

Corollary 5 *Suppose that* $A_1, A_2, \ldots, A_n$ *are subsets of a set* X*, and suppose that, for some fixed* r*, we have*

(a) $|A_i| \geq r$ *for all* $1 \leq i \leq n$*, and*

(b) each $x \in X$ *belongs to* at most r *subsets* A_i*.*

Then there is a SDR for the A_i.

Proof: We need to show that Hall's condition is satisfied. So pick any subset $I \subseteq \{1, 2, \ldots, n\}$ and try to show that

$$\left| \bigcup_{i \in I} A_i \right| \geq |I|.$$

Let

$$P = \{(x, A_i) \mid x \in A_i, i \in I\},$$

and estimate the size of P in two different ways.

First estimate:

Stage 1: Choose $x \in \bigcup_{i \in I} A_i$. Number of choices is just $|\bigcup_{i \in I} A_i|$.

Stage 2: For this particular x, choose $i \in I$ such that $x \in A_i$. By condition (b) there are *at most* r choices.

Hence $|P| \leq r|\bigcup_{i \in I} A_i|$.

Second estimate:

Stage 1: Choose A_i with $i \in I$. Number of choices is $|I|$.

Stage 2: Choose $x \in A_i$. Number of choices is $|A_i| \geq r$ by condition (a).

Hence $|P| \geq r|I|$.

Putting these two inequalities together we get

$$r \left| \bigcup_{i \in I} A_i \right| \geq |P| \geq r|I|,$$

and dividing through by r we get, as required,

$$\left| \bigcup_{i \in I} A_i \right| \geq |I|.$$

So the A_i satisfy Hall's condition, so by Hall's theorem there is an SDR.

We might also be interested in *how many* SDRs there are, in the case when Hall's condition is satisfied. Of course, this depends on how many elements each set A_i has: if $|A_i| = 1$, then obviously there is only one possible representative.

Lecture 21, 24/11/11

Theorem 12 *Suppose that the sets $A_1, A_2, \ldots, A_n$ satisfy Hall's condition (so*

$$\left| \bigcup_{i \in I} A_i \right| \geq |I|$$

for all $I \subseteq \{1, 2, \ldots, n\}$), and $|A_i| \geq r$ for all i.

(a) If $n \geq r$, then the number of SDRs is at least $r!$.

(b) If $n < r$, then the number of SDRs is at least $(r)_n = r(r-1).\cdots(r-n+1)$.

Proof: we modify the proof of Hall's theorem above. In the base case of the induction, $n = 1$, and the number of SDRs is at least r. This is equal to $r!$ in the case $r \leq n = 1$, and to $(r)_n$ in the other case.

Now consider the inductive step. If we are in the situation where there is no critical set, then we can choose an arbitrary element $a_n \in A_n$ as the representative (in at least r ways), and complete to an SDR, by taking any SDR of the sets $A'_1, \ldots, A'_{n-1}$, where $A'_i = A_i \setminus \{a_n\}$. Now $|A'_i| \geq r-1$, and there are $n-1$ sets A'_i, so the number of such SDRs is (a) at least $(r-1)!$ in the case $n-1 \geq r-1$, or (b) at least $(r-1)_n$ in the case $n-1 < r-1$. Hence the number of SDRs of $A_1, \ldots, A_n$ is at least r times this, as required.

If on the other hand there is a critical set J, of size m, say, then the proof divides into two cases according to whether $m \geq r$ or $m < r$. If $m \geq r$, then by induction the sets A_j, $j \in J$ already have at least $r!$ SDRs, each of which extends to at least one SDR of the whole collection of A_i. The case $m < r$ was unfortunately omitted from the lecture as given: in this case we only know that the number of SDRs of the A_j, $j \in J$, is at least $r(r-1).\cdots(r-m+1)$, so we need to estimate the number of ways of extending each of them to a SDR of the whole collection.

In this case, we constructed sets $A_i^* = A_i \setminus \bigcup_{j \in J} A_j$, by removing at most m elements from each A_i, so that $|A_i^*| \geq r-m$. Now the number of sets A_i^* is $n-m$, so in case (a), we have $n-m \geq r-m$ and there are at least $(r-m)!$ ways of choosing the SDR of the A_i^*. Hence there are at least $(r)_m(r-m)! = r!$ SDRs for the whole collection of the A_i. Similarly, in case (b), we have $n-m < r-m$ and there are at least $(r-m)_{n-m}$ choices for the SDR of the A_i^*, making at least $(r)_m(r-m)_{n-m} = (r)_n$ altogether.

Chapter 6

Latin squares

6.1 Counting Latin squares

Definition 9 *A* Latin square of order n *is an* $n \times n$ *matrix whose entries come from a set of* n *symbols (usually* $\{1, 2, \ldots, n\}$*, or* $\{0, 1, 2, \ldots, n-1\}$*), such that each symbol appears exactly once in each row, and exactly once in each column.*

Example 51

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}.$$

Is it obvious from this that we can construct a Latin square of order n for any n?

If not, let the first row be $(1, 2, 3, \ldots, n)$, the second row $(2, 3, 4, \ldots, n, 1)$ and so on: keep rotating the coordinates backwards one step at a time.

Example 52 *A completed sudoku is an example of a* 9×9 *Latin square (with extra properties).*

We will try to address the question of how many Latin squares of order n there are. You probably appreciate that there are a large number of sudoku squares, and so there are likely to be a large number of Latin squares also. But how large?

First we use Hall's Marriage Theorem to show that we can construct a Latin square one row at a time: it is not possible to get stuck, provided we construct a whole row at a time.

Definition 10 *A* $k \times n$ *Latin rectangle is a* $k \times n$ *matrix, where* $k \leq n$*, with entries from a set of* n *elements (usually* $\{1, 2, 3, \ldots, n\}$*), such that no symbol appears twice in any row or column.*

Thus an $n \times n$ Latin rectangle is the same thing as a Latin square of order n.

Example 53

$$M = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \\ 4 & 1 & 5 & 3 & 2 \end{pmatrix}$$

is a 3×5 *Latin rectangle.*

Lemma 3 *If* $k < n$*, then any* $k \times n$ *Latin rectangle can be extended to a* $(k+1) \times n$ *Latin rectangle.*

First consider our example, and, for $1 \leq i \leq 5$, let A_i be the set of options for the ith entry in the next row. Then we have $A_1 = \{2, 5\}$, $A_2 = \{3, 4\}$, $A_3 = \{2, 4\}$, $A_4 = \{1, 5\}$, $A_5 = \{1, 3\}$. We need a *system of distinct representatives* for these sets.

Now we see that $|A_i| = 2 \geq 1$, and $|A_i \cup A_j| \geq 2$, and $|A_i \cup A_j \cup A_k| \geq 3$, and so on. Hence Hall's condidion is satisfied, and there is a SDR, say

$$\{(2, A_1), (3, A_2), (4, A_3), (5, A_4), (1, A_5)\}.$$

So the corresponding 4×5 Latin rectangle is

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \\ 4 & 1 & 5 & 3 & 2 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}.$$

Proof of Lemma: (just a generalisation of the argument in this special case). For $1 \leq i \leq n$, let A_i be the set of symbols that have not been used so far in the ith column. Then an SDR for these sets gives us a possible $(k+1)$st row for the Latin rectangle.

First, we have $|A_i| = n - k$. Now choose any one of the n symbols, and call it j. Now j occurs exactly once in each row of the $k \times n$ rectangle, so occurs exactly k times, and therefore occurs in exactly k columns. Hence j lies in exactly the other $n - k$ sets A_i.

But we proved a corollary to Hall's theorem dealing with this case: $|A_i| \geq n-k$ and no element occurs in more than $n-k$ of the $|A_i|$. This corollary implies that there exists a SDR. This proves the lemma.

Lecture 22, 28/11/11

The first half of Lecture 22 was proving the Theorem at the end of Chapter 5, because both the statement and the proof given in Lecture 21 were incorrect. Apologies for the confusion caused.

Theorem 13 *Any* $k \times n$ *Latin rectangle can be completed to a Latin square of order* n*.*

Proof: The lemma guarantees that we can keep adding rows one at a time until the rectangle becomes square.

We could also ask, at each stage, *how many* ways are there of adding the next row? In other words, how many SDRs are there for the corresponding sets A_i?

Suppose we have a $k \times n$ Latin rectangle, and as before let A_i be the set of symbols which have not already been used in the ith column. Then $|A_i| = n - k$, and we already showed that the A_i satisfy Hall's condition. Hence the theorem from the end of the last chapter implies that there are at least $(n-k)!$ SDRs. That is, there are at least $(n-k)!$ choices for the next row of the Latin square.

So, there are $n!$ choices for the first row of the Latin square, and at least $(n-1)!$ choices for the second row, and at least $(n-2)!$ choices for the third row, and so on. Hence the total number of Latin squares of order n is at least

$$n!(n-1)!(n-2)!\cdots 3!2!1! = \prod_{k=1}^{n} k!$$

Examples: if $n = 3$, this gives $3!2! = 12$, which is in fact the total number of Latin squares of order 3. If $n = 4$, it gives $4!3!2! = 288$. However, there are in fact 576 Latin squares of order 4 (exercise) over any given set of four symbols. The numbers of Latin squares of order up to 11 are known exactly.

So far we have seen a *lower* bound for the number of Latin squares. What about an upper bound?

Theorem 14 *The number of Latin squares of order n is at most $(n!)^n/e^{n-1}$ (provided $n \geq 3$).*

Proof: There are $n!$ choices for the first row, and every other row is a derangement of the first. Now the number of derangements is approximately $n!/e$, and if we use this figure for all the remaining rows, we get a total number of possibilities

$$n!\left(\frac{n!}{e}\right)^{n-1} = \frac{n!^n}{e^{n-1}}.$$

This is not quite correct, however, as the number of choices for the second row is *more* than $n!/e$, if n is even. Lecture 23, 29/11/11

To correct this, observe that the second and third rows are different, so the number of choices for the two rows is at most $k(k-1)$, where (in the case n even) k is the integer satisfying

$$k - \frac{1}{2} \leq \frac{n!}{e} \leq k.$$

Hence $k(k-1) < (k-\frac{1}{2})^2 \leq (n!/e)^2$. Now for each of the subsequent rows, there are at most $k - 2 < n!/e$ choices. So the result follows.

Notice that in fact the bound does not hold for $n = 2$: there are two Latin squares in this case, but $2!^2/e < 2$.

In fact, this bound is not very good, as the third row has to be a derangement of both the first two rows, etc, so the number of choices at each stage is a good deal less than this estimate. But the number of choices at each stage is *not* independent of previous choices, so there is no easy way of counting Latin squares exactly.

6.2 Mutually orthogonal Latin squares

Definition 11 *Suppose that $A = (a_{ij})$ and $B = (b_{ij})$ are two Latin squares of order n. Then A and B are called (mutually)* orthogonal *if the n^2 pairs (a_{ij}, b_{ij}) are all different.*

In other words, all possible pairs occur exactly once. (Note: 'mutually orthogonal' means 'orthogonal to each other'. We abbreviate the phrase 'mutually orthogonal Latin squares' to MOLS, for convenience.)

Example 54 *The Latin squares*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}$$

are (mutually) orthogonal, because all 9 pairs occur:

$$\begin{pmatrix} (1,1) & (2,2) & (3,3) \\ (2,3) & (3,1) & (1,2) \\ (3,2) & (1,3) & (2,1) \end{pmatrix}.$$

Example 55 *There are only two Latin squares of order 2, and they are not orthogonal:*

$$\begin{pmatrix} (1,2) & (2,1) \\ (2,1) & (1,2) \end{pmatrix}$$

has the pais $(1,2)$ and $(2,1)$ twice each, and does not have $(1,1)$ or $(2,2)$.

Problem: for given n, how many mutually orthogonal Latin squares can there be? (Answer not known in general.)

Example 56 *For $n = 2$ we cannot have two MOLS, so the maximum number is 1.*

Example 57 *For $n = 3$ we have two MOLS. We cannot have any more, because relabelling the symbols in two Latin squares does not change whether they are orthogonal or not, so we may assume all our MOLS have first row $(1,2,3)$. But then there only are two possible Latin squares.*

So we might conjecture that the maximum number of Latin squares of order n is $n-1$. In fact, it is *at most* $n-1$, but it is a long-standing open problem to determine for which n it is *actually possible* to find $n-1$ MOLS.

Theorem 15 *It is not possible to find more than $n-1$ MOLS of order n.* Lecture 24, 1/12/11

Proof: Suppose $A_1, \ldots, A_n$ are n MOLS of order n. We aim for a contradiction. First, we may assume that every one of $A_1, \ldots, A_n$ has $(1, 2, \ldots, n)$ as its first row. Thus, when we compare A_i and A_j we see all the pairs (x, x) accounted for by the first row. In particular, the pair $(1, 1)$ is already accounted for.

Now where can we put the symbol 1 in the second row of the A_i? Well, it can't go in the first column, so it has to go in one of the remaining $n-1$ columns. Could we put the symbol 1 in the same place in A_i as in A_j? No, because then we'd have the pair $(1, 1)$ twice when combining A_i with A_j, contradicting the assumption that A_i and A_j are mutually orthogonal.

Hence the symbol 1 has to go in n different places in $A_1, \ldots, A_n$. But there are only $n-1$ places to choose from. Contradiction.

Finite field construction for MOLS. You know already that if p is prime, then the integers modulo p form a *field*. In particular, there is a finite field of order p. (The order of a field is just the number of elements in it.)

In fact, there is exactly one finite field of each order p^d, where p is prime and d is a positive integer, and no finite fields of any other order. However, if you prefer just to think of the fields of prime order, that is, the integers modulo p, then you will not lose much.

Example 58 *Our two MOLS of order 3 can be re-written with the symbols $0, 1, 2$, understood as integers modulo 3, as follows:*

$$A = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix}.$$

What does A remind you of, for mod 3 arithmetic? It is the addition table! Thus the entry (a_{ij}) in the ith row and jth column is $i + j$, that is $a_{ij} = i + j$.

What about B? Is there a nice formula for b_{ij}? Well, yes, it is $b_{ij} = 2i + j$.

The general construction: let F be a finite field, and let $f \in F \setminus \{0\}$. Then for each f, define the matrix $A_f = (a_{ij}^{(f)})$, with rows and columns indexed by $i, j \in F$, by

$$a_{ij}^{(f)} = if + j.$$

Now I claim that the A_f are mutually orthogonal Latin squares.

Theorem 16 *With the above notation, and for any ordering of the elements of the field F,*

(a) for any $f \neq 0$, the matrix A_f is a Latin square;

(b) if $f \neq g$, then the Latin squares A_f and A_g are mutually orthogonal.

Lecture 25, 5/12/11

Proof of (a): two entries in the ith row are $if + j_1$ and $if + j_2$, and if these are equal, $if + j_1 = if + j_2$, so $j_1 = j_2$, in other words they are the same entry.

Similarly, two entries in the jth column are $i_1 f + j$ and $i_2 f + j$. If these are equal, then $i_1 f = i_2 f$, and because $f \neq 0$ we can divide by f to get $i_1 = i_2$, which again means the two entries are in the same row and column, so in the same place.

Proof of (b): taking the (i, j) entries from the two Latin squares, we have the ordered pair $(if + j, ig + j)$. We want to show that as i and j vary, we get all n^2 pairs, where n is the order of the field. In other words, we want to show that we cannot get any pair more than once.

So suppose that we get the same pair in the entry in the position (i, j) and in the entry in the position (x, y). Then $(if + j, ig + j) = (xf + y, xg + y)$. This means

$$\begin{aligned} if + j &= xf + y \\ ig + j &= xg + y \end{aligned}$$

Subtracting these equations gives $i(f - g) = x(f - g)$, and if $f \neq g$ we can divide by $f - g$ to get $i = x$. Substituting back into the first equation then gives $j = y$. In other words the pair $(i, j) = (x, y)$ as required.

Example 59 *Three MOLS of order 4. There is a field of order 4 whose elements I shall write as $0, 1, a, b$, where $b = a + 1$, and $a^2 = b$. The addition and multiplication tables are as follows:*

$+$	0	1	a	b
0	0	1	a	b
1	1	0	b	a
a	a	b	0	1
b	b	a	1	0

$\times$	0	1	a	b
0	0	0	0	0
1	0	1	a	b
a	0	a	b	1
b	0	b	1	a

The three Latin squares described by the above construction are

$$A_1 = \begin{pmatrix} 0 & 1 & a & b \\ 1 & 0 & b & a \\ a & b & 0 & 1 \\ b & a & 1 & 0 \end{pmatrix}, A_a = \begin{pmatrix} 0 & 1 & a & b \\ a & b & 0 & 1 \\ b & a & 1 & 0 \\ 1 & 0 & b & a \end{pmatrix}, A_b = \begin{pmatrix} 0 & 1 & a & b \\ b & a & 1 & 0 \\ 1 & 0 & b & a \\ a & b & 0 & 1 \end{pmatrix}.$$

You can easily check that any two are mutually orthogonal.

This theorem implies that if n is a power of a prime, then there exist $n-1$ MOLS of order n.

On the other hand, it is known that there do not exist two MOLS of order 6. This was an old problem due to Euler (18th century), not solved until the 20th century.

It is an open problem whether or not there exists any other value of n for which there exist $n-1$ MOLS.

Product construction for MOLS. Suppose A is a Latin square of order n and B is a Latin square of order m, then we can construct the *product* Latin square $A \circ B$ as follows.

Suppose the symbols in A are $1, 2, 3, \ldots, n$. We take n copies of B, each with a different set of symbols. Call these Latin squares $B_1, \ldots, B_n$. Then replace each symbol i in A by the square B_i. In this way we obtain an $nm \times nm$ square.

Example 60 *Take*

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix},$$

and then we can take two copies of B as

$$B_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}, B_2 = \begin{pmatrix} a & b & c \\ b & c & a \\ c & a & b \end{pmatrix}.$$

Then

$$A \circ B = \begin{pmatrix} B_1 & B_2 \\ B_2 & B_1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & a & b & c \\ 2 & 3 & 1 & b & c & a \\ 3 & 1 & 2 & c & a & b \\ a & b & c & 1 & 2 & 3 \\ b & c & a & 2 & 3 & 1 \\ c & a & b & 3 & 1 & 2 \end{pmatrix}.$$

It is easy to see that $A \circ B$ is a Latin square: the same symbol cannot occur in different blocks in the same row or column of block; and each block is a Latin square, so within one block the same symbol cannot occur twice in the same row or column.

Lemma 4 *If A and B are MOLS of order n, and C and D are MOLS of order m, then $A \circ C$ and $B \circ D$ are MOLS of order mn.*

Remark: $A \circ C$ is *not* orthogonal to $A \circ D$ or to $B \circ C$. Why not?

Lecture 26, 6/12/11

Example 61 *Take*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}$$

and

$$C = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

Then

$$A \circ C = \begin{pmatrix} C_1 & C_2 & C_3 \\ C_2 & C_3 & C_1 \\ C_3 & C_1 & C_2 \end{pmatrix}, B \circ D = \begin{pmatrix} D_1 & D_2 & D_3 \\ D_3 & D_1 & D_2 \\ D_2 & D_3 & D_1 \end{pmatrix}.$$

Now suppose that some pair of symbols (c, d) occurs twice in the pair $(A \circ C, B \circ D)$. Then the symbol c has to occur in a particular one of the C_i, say C_1. Similarly, the symbol d occurs in just one of the D_i, say D_1. So (c, d) can only occur in one of the blocks, in this case the (C_1, D_1) block, i.e. the top left hand corner. This is because A and B are MOLS.

Finally, in a particular block (C_i, D_j) the pair (c, d) occurs exactly once, since C and D are MOLS.

Corollary 6 *Suppose we have x MOLS of order n and y MOLS of order m. Then we can construct* $\min\{x, y\}$ *MOLS of order mn.*

Corollary 7 *If n is a positive integer and $n = p_1^{m_1}.p_2^{m_2}.\cdots.p_r^{m_r}$, where the p_i are distinct primes and*

$$p_1^{m_1} < p_2^{m_2} < \cdots < p_r^{m_r},$$

then there are $p_1^{m_1}$ MOLS of order n.

This is because there are (at least) $p_1^{m_1} - 1$ MOLS of each of the orders $p_1^{m_1}$, $p_2^{m_2}$, ... (because there are fields of these orders). Then the product construction gives (by induction on r) $p_1^{m_1}$ MOLS of order n.

Solutions to Exercises 8. Q.1 should have been easy: in (a) you find four sets with only the three elements $1, 3, 5$ between them, so Hall's condition is not satisfied. In (b) trial and error is sufficient to find a SDR; but you can also use the proof of Hall's theorem, because there is a critical set $\{1, 3, 5\}$, and therefore you might as well choose a SDR for this sub-collection first. Then remove $1, 3, 5$ from all the remaining sets, and choose an SDR for them.

Q.2 (a) You are asked to choose three sets $A_1, A_2, A_3 \subseteq \{1, 2, 3\}$ (not necessarily distinct!) which have various numbers of SDRs. For example, if you choose

$A_1 = \{1\}$, $A_2 = \{2\}$, $A_3 = \{3\}$, then there is only one SDR. At the other extreme, if you choose $A_1 = A_2 = A_3 = \{1, 2, 3\}$, then you can pick any representatives you like, so in $3! = 6$ ways.

(b) Now if 5 of these 6 ways of picking an SDR are possible, then you still need all of $1, 2, 3 \in A_1$, and similarly for A_2 and A_3. But then the sixth SDR is also possible.

Q.3 Proof by induction on n. If there is no critical set (as in the proof of Hall's theorem), then the proof shows that we can pick any $a_n \in A_n$ and extend to an SDR. On the other hand, if there is a critical set J, say, we can pick an SDR of the A_j for $j \in J$ first: and, by induction, there is some A_j for which any $a_j \in A_j$ will do.

Q.4 Follow the hint, and let $A_i' = A_i \cup \{z_1, \ldots, z_m\}$. Then the A_i' satisfy Hall's condition, so have an SDR. At most m of the A_i' have one of the z_j as their representative, so if we remove these, we have at least $n - m$ sets A_i' which have elements from the original A_i as representatives.

Chapter 7

Extremal set theory

Lecture 27, 8/12/11

The typical problem in extremal set theory is, what is the maximum number of sets (of a certain type) that can be chosen with certain properties? The properties we consider are not properties of *individual* sets, as these problems are just the ordinary counting problems we considered at the beginning of this course. Instead, we now consider properties of *relations between sets*. For example, we might want

- all intersections to be empty;
- all intersections to be non-empty;
- all intersections to have size 1;
- no set to be a subset of any other;

or any of a myriad other possible conditions.

The second type of problem in extremal set theory is, once we know the maximum number of sets as above, what are all the configurations of sets which reach this maximum? For example, are they all essentially the same?

In this chapter we will always be considering subsets of a set X, and some unknown family $\mathcal{F}$ of (distinct) subsets of X. Thus $\mathcal{F} \subseteq \mathcal{P}(X)$, the power set of X, where $\mathcal{P}(X) = \{Y \mid Y \subseteq X\}$.

7.1 Intersecting families

Definition 12 *If X is a set and $\mathcal{F}$ is a family of (distinct) subsets of X, then $\mathcal{F}$ is an* intersecting family *if, for every choice of $A, B \in \mathcal{F}$, we have $A \cap B \neq \emptyset$.*

First of all, how large do we think an intersecting family $\mathcal{F}$ can be, if $|X| = n$? After a bit of thought, you will probably realise that to get non-empty intersections, we can choose a particular element $a \in X$, and take all the sets containing a. Then every intersection contains a, so is non-empty.

More formally, choose $a \in X$, and let $\mathcal{F} = \{Y \subseteq X \mid a \in Y\}$. Then if $A, B \in \mathcal{F}$, we have $a \in A$ and $a \in B$, so $a \in A \cap B$, and therefore $A \cap B \neq \emptyset$.

How big is this family $\mathcal{F}$? Well, exactly half of the subsets of X contain the element a, so $|\mathcal{F}| = 2^{n-1}$.

Can we do better than this, or is this family extremal?

Theorem 17 *If $|X| = n$ and $\mathcal{F}$ is an intersecting family of subsets of X, then $|\mathcal{F}| \leq 2^{n-1}$.*

Proof: We know that there are exactly 2^n subsets of X. That is $|\mathcal{P}(X)| = 2^n$. We partition $\mathcal{P}(X)$ into sets of size 2, by pairing each subset of X with its complement. Formally, the partition is

$$\mathcal{Q} = \{\{A, A^c\} \mid A \subseteq X\}$$

where the complement of A is $A^c = X \setminus A = \{x \in X \mid x \notin A\}$.

Now if $A \in \mathcal{F}$, then $A^c \notin \mathcal{F}$, because $A \cap A^c = \emptyset$. So when we choose sets to put in $\mathcal{F}$, we can take at most one of each pair $\{A, A^c\}$. Since there are 2^{n-1} such pairs, we have $|\mathcal{F}| \leq 2^{n-1}$.

Example 62 *If $n = 3$, say $X = \{1, 2, 3\}$, so that*

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Then the partition $\mathcal{Q}$ of $\mathcal{P}(X)$ is into four parts: $\mathcal{Q} = \{P_1, P_2, P_3, P_4\}$ where

$$P_1 = \{\emptyset, \{1, 2, 3\}\}, P_2 = \{\{1\}, \{2, 3\}\}, P_3 = \{\{2\}, \{1, 3\}\}, P_4 = \{\{3\}, \{1, 2\}\}.$$

That is

$$\mathcal{Q} = \{\{\emptyset, \{1, 2, 3\}\}, \{\{1\}, \{2, 3\}\}, \{\{2\}, \{1, 3\}\}, \{\{3\}, \{1, 2\}\}\}.$$

So $|\mathcal{Q}| = 4 = 2^{n-1}$, and any intersecting familiy can contain at most one of the two sets in P_i, for each i.

Let us look at this example in more detail, and see what kind of extremal families we can find. First, what element of P_1 can we take? It cannot be $\emptyset$, so it must be $\{1, 2, 3\}$. Now look at P_2. If we choose $\{1\}$ to be in our family, then every other set in our family has to intersect this, so contains 1. Hence our family must be

$$\mathcal{F}_1 = \{\{1, 2, 3\}, \{1\}, \{1, 3\}, \{1, 2\}\}.$$

Could we choose the other possibility, $\{2, 3\}$? Yes (for example by symmetry). But could we choose *none* of the sets of size 1? If so, we would have

$$\mathcal{F}_2 = \{\{1, 2, 3\}, \{2, 3\}, \{1, 3\}, \{1, 2\}\}.$$

Does this work? Yes!

So, at least in this example, we have two different types of extremal families. The first type, like $\mathcal{F}_1$, consists of all sets containing a particular element, say 1. How would we describe the second type, $\mathcal{F}_2$? It consists of all the 'big' sets, that is, subsets A of X with $|A| > \frac{1}{2}|X|$.

Does this generalise to all X? Well, it works OK if n is odd. Then if A and B both have more than $n/2$ elements, they cannot be disjoint (for then X would have more than n elements). From each pair $\{A, A^c\}$ we always choose the bigger one. So we get an intersecting family of 2^{n-1} sets in this way.

But what about the case when n is even? Then we have some cases where $|A| = |A^c| = n/2$. In fact we can make arbitrary choices in these cases, and still get an intersecting familiy. (Why?)

So there are two different kinds of extremal families. Are there any more? (This is a difficult problem! In general there may be many different types.)

7.2 Sperner families

Lecture 28, 12/12/11

Definition 13 *A* Sperner family *of subsets of a set X is a family $\mathcal{F} \subseteq \mathcal{P}(X)$ with the property that if $A, B \in \mathcal{F}$ than $A \not\subseteq B$.*

Example 63 *If $X = \{1, 2, 3, 4\}$ then the family*

$$\mathcal{F} = \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}\}$$

of all subsets of size 2 is a Sperner family.

Indeed, it is the (unique) largest possible Sperner family in this instance. We shall show that the largest possible Sperner family of a set of n elements has size

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

Moreover, we shall completely classify the Sperner families of this maximal size.

First we show that such families exist.

Example 64 *If $|X| = n$, then the set of all subsets of X with size $\lfloor \frac{n}{2} \rfloor$ is a Sperner family.*

Proof: if A and B are two distinct subsets of the same size, then obviously $A \not\subseteq B$. The size of this family is clearly

$$\binom{n}{\lfloor \frac{n}{2} \rfloor}.$$

Proving that there is no bigger Sperner family is a little technical.

Lemma 5 *Let $n \geq 1$ be a fixed integer. Then $\binom{n}{k}$ takes its maximum value when $k = n/2$ (if n is even), or $k = (n \pm 1)/2$ (if n is odd).*

Proof: if $k < n/2$ then

$$\binom{n}{k+1} / \binom{n}{k} = \frac{n-k}{k+1} > 1.$$

Then use $\binom{n}{n-k} = \binom{n}{k}$ to prove a similar result for $k > n/2$.

The main idea of the proof of Sperner's theorem is to count the (maximal) *chains* of subsets, that is, chains

$$\emptyset \subset Y_1 \subset Y_2 \subset \cdots Y_{n-1} \subset Y_n = X$$

where $|Y_i| = i$ for all i. Since we have to choose one of the n elements to be in Y_1, then one of the remaining $n-1$ to be in Y_2, and so on, the total number of chains for X is $n!$.

On the other hand, if we specify that Y_k is a fixed set A, say, where $|A| = k$, then the first k elements have to be chosen (in order) from the elements of A, that is in $k!$ ways. Then the remaining $n-k$ elements are chosen in order in $(n-k)!$ ways. Thus the number of chains that go through A is $k!(n-k)! = |A|!(n-|A|)!$.

Now if A and B lie in the same chain, then one of them is a subset of the other. Thus if A and B come from a Sperner family, they cannot both lie in the same chain. This gives us another way of counting chains, by dividing them up according to which element (if any) of the Sperner family they go through.

More precisely, if $\mathcal{F} = \{A_1, A_2, \ldots, A_t\}$ is a Sperner family, then the total number of chains is no more than the sum over all A_i, of the number of chains through A_i, that is

$$n! \geq \sum_{i=1}^{t} |A_i|!(n-|A_i|)!$$

Dividing through by $n!$ gives

$$\begin{aligned} 1 &\geq \sum_{i=1}^{t} \frac{1}{\binom{n}{|A_i|}} \\ &\geq \sum_{i=1}^{t} \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \\ &= \frac{t}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \end{aligned}$$

which gives $\binom{n}{\lfloor \frac{n}{2} \rfloor} \geq t$ as required.

Moreover, if t is actually equal to this number, then we must have equality all the way through the proof as well. In particular $\binom{n}{|A_i|} = \binom{n}{\lfloor \frac{n}{2} \rfloor}$ for all i. Therefore,

in the case when n is even, all the sets A_i have $|A_i| = n/2$. In the case when n is odd, we have $|A_i| = (n \pm 1)/2$. In fact, it is not too hard to show that in this latter case, all the A_i have the same size, that is you cannot have sets of both sizes $(n \pm 1)/2$ in the same extremal Sperner family.

7.3 The Erdős–de Bruijn Theorem

We shall only state this theorem, not prove it. (A proof can be found in Prof. Cameron's notes.) Lecture 29, 13/12/11

Theorem 18 *If $A_1, A_2, \ldots, A_b$ are distinct subsets of $X = \{1, 2, \ldots, n\}$ such that $|A_i \cap A_j| = 1$ for all $i \neq j$, then $b \leq n$. Moreover, if $b = n$, then (after re-labelling if necessary) one of the following holds:*

(a) $A_1 = \{1, n\}, A_2 = \{2, n\}, \ldots, A_{n-1} = \{n-1, n\}$, $A_n = \{n\}$;

(b) $A_1 = \{1, n\}, A_2 = \{2, n\}, \ldots, A_{n-1} = \{n-1, n\}$, $A_n = \{1, 2, \ldots, n-1\}$;

(c) every pair of elements of X lies in exactly one of the sets A_i.

In case (c), we might as well assume that there is a set of four elements of X with the property that no three lie in the same A_i: for otherwise we are in case (b). With this extra assumption, case (c) is known as a *projective plane*. Usually, in this case, the elements of X are called *points*, and the sets A_i are called *lines*.

Definition 14 *A* projective plane *is a collection of points and lines with the properties:*

- *every pair of lines intersects in exactly one point;*
- *every pair of points lie on exactly one line;*
- *there exist four points such that no three of them lie on a line.*

Example 65 *If $X = \{1, 2, 3, 4, 5, 6, 7\}$ then the family*

$$\mathcal{F} = \{\{1, 2, 3\}, \{1, 4, 5\}, \{1, 6, 7\}, \{2, 4, 6\}, \{2, 5, 7\}, \{3, 4, 7\}, \{3, 5, 6\}\}$$

is a projective plane.

We see in this example that every line has the same number of points, namely 3, and moreover every point lies on 3 lines. The analogous properties hold in general: there is an integer n, called the *order* of the projective plane, such that every line has $n + 1$ points, every line lies on $n + 1$ lines, and the total number of points is $n^2 + n + 1$ (and therefore the total number of lines is also $n^2 + n + 1$).

To prove this, suppose that the points $1, 2, 3, 4$ have the property that no three lie on a line. In particular, 1 does not lie on the line through $3, 4$. Hence the number of lines through 1 is equal to the number points on the line $3-4$ (for every line through one meets the line $3-4$ somewhere, so just join 1 to each of the points on this line in turn: these must be distinct lines). By symmetry, this is also equal to the number of lines through 2, and to the number of points on $2-3$, and $2-4$, and so on.

Now pick any point P other than $1, 2, 3, 4$: it cannot lie on two of the lines $2-3$, $2-4$ and $3-4$, because their intersections are already accounted for, so by joining it to the points of one of these lines it is not on, we get that the number of lines through P is the same as the number of lines through 1.

Similarly, any line L must miss at least two of the points $1, 2, 3, 4$, so by joining its points to one of these we see that it also has the same number of points on it as does $3-4$.

7.4 Projective planes and MOLS

Lecture 30, 15/12/11

We shall show that existence of a projective plane of order n is equivalent to existence of $n-1$ MOLS of order n. But the projective plane typically shows more symmetry. Indeed, to go from a projective plane to a set of $n-1$ MOLS one first has to choose two points in the projective plane: it is possible that a different choice of points will give a completely different set of MOLS.

So, pick two points x, y in our projective plane of order n. There are $n-2$ further points on the line xy, and n^2 other points. Each of these n^2 points is on one of the n lines through x but not through y; and is on one of the n lines through y but not through x. Moreover each of the n lines through x meets each of the n lines through y in exactly one of the n^2 points.

This groups our n^2 points into n *rows* (i.e. the lines through x) and n *columns* (i.e. the lines through y). Next we construct some Latin squares on these rows and columns.

Pick any other point z on the line xy, and consider the n lines through z but not through x or y. Each such line intersects every row and every column in exactly one point. Label these n points by a symbol identifying this line. Do the same with $n-1$ more symbols for the other $n-1$ lines through z. This gives us a Latin square on the rows and columns.

Now do the same thing for another point t on the line xyz. We get another Latin square. So each point is labelled by a pair consisting of a line through z and a line through t. We cannot have the same label at two different points, for then that pair of lines would intersect in two points. Hence the two Latin squares are orthogonal to each other.

Thus, if we do this for all $n-1$ of the points on xy (other than x, y themselves) we get $n-1$ mutually orthogonal Latin squares.

Conversely, given $n-1$ MOLS of order n, we can reverse this process to get a projective plane of order n.