Math 222 Fall 2022, Lecture 11: Introduction

website: `https://www.cip.ifi.lmu.de/~grinberg/t/22fco`

1. Introduction (cont'd)

1.5. Counting tuples and maps (cont'd)

Recall: In Lecture 10, we proved the formula

$$|A^n| = |A|^n \tag{1}$$

for any finite set A and any $n \in \mathbb{N}$.

1.5.2. Counting maps

Definition 1.5.5. Let A and B be two sets. Then, B^A shall mean the set of all maps (= functions) from A to B.

This strange notation is used for a reason:

Theorem 1.5.6. Let A and B be two finite sets. Then, the set B^A is again finite, and has size

$$\left|B^A\right| = |B|^{|A|}.$$

Example 1.5.7. Let us see how this works out for $A = [2] = \{1, 2\}$ and $B = [3] = \{1, 2, 3\}$. Then,

$$B^A = \{\text{all maps from } A \text{ to } B\} = \{\text{all maps from } [2] \text{ to } [3]\}.$$

One way to list these maps is by using **two-line notation**. The **two-line notation** of a map $f : A \to B$ is a table with two rows, where the top row lists all elements of A, and the bottom row lists their images under f. In other words, it is a table of the form

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ f(a_1) & f(a_2) & \cdots & f(a_k) \end{pmatrix},$$

where $a_1, a_2, \ldots, a_k$ are the elements of A. The maps from $A = [2]$ to $B = [3]$ are then (in two-line notation)

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

There are 9 of these maps, which fits the claim of the theorem, since $|B|^{|A|} = 3^2 = 9$.

Proof of Theorem 1.5.6 (sketched). Let $a_1, a_2, \ldots, a_k$ be all the elements of A, listed in some way without repetition. Then, $k = |A|$.

Recall that B^A is the set of all maps $f : A \to B$. Such a map $f : A \to B$ is uniquely determined by the values $f(a_1), f(a_2), \ldots, f(a_k)$, and these values can be chosen arbitrarily (and independently) from B. So essentially, such a map f is "the same as" the k-tuple $(f(a_1), f(a_2), \ldots, f(a_k)) \in B^k$. Formally speaking, there is a bijection

$$B^A \to B^k,$$
$$f \mapsto (f(a_1), f(a_2), \ldots, f(a_k)).$$

Thus, by the bijection principle,

$$\left|B^A\right| = \left|B^k\right| = |B|^k \qquad \text{(by (1))}$$
$$= |B|^{|A|} \qquad (\text{since } k = |A|).$$

This proves Theorem 1.5.6. □

1.5.3. Application: Counting subsets bijectively

Recall:

Theorem 1.5.8. Let S be an n-element set (where $n \in \mathbb{N}$). Then,

$$(\# \text{ of subsets of } S) = 2^n.$$

We proved this by induction in §1.4.1 (Lecture 8). Now, let us reprove this bijectively:

Another proof of Theorem 1.5.8 (sketched). (See §1.5.3 in the 2019 notes for details.)

Let $\mathcal{P}(S)$ denote the powerset of S; this is the set of all subsets of S. Thus, we must prove that $|\mathcal{P}(S)| = 2^n$.

Let $s_1, s_2, \ldots, s_n$ be the n elements of S, listed in some way (without repetitions).

We will construct a bijection

$$\text{from the set } \mathcal{P}(S) = \{\text{subsets of } S\}$$
$$\text{to the set } \{0,1\}^n = \{n\text{-tuples of elements of } \{0,1\}\}.$$

To wit, we define a map $A : \mathcal{P}(S) \to \{0,1\}^n$ as follows:

$$A(T) = ([s_1 \in T],\ [s_2 \in T],\ \ldots,\ [s_n \in T]) \qquad \text{for any } T \subseteq S.$$

Conversely, we define a map $B : \{0,1\}^n \to \mathcal{P}(S)$ as follows:

$$B(i_1, i_2, \ldots, i_n) = \{s_k \mid k \in [n] \text{ and } i_k = 1\} \qquad \text{for each } (i_1, i_2, \ldots, i_n) \in \{0,1\}^n .$$

It is straightforward to see that

$$A \circ B = \operatorname{id} \qquad \text{and} \qquad B \circ A = \operatorname{id} .$$

(The notation id stands for the identity map of a set, i.e., the map that sends each x to x. So "$A \circ B = \operatorname{id}$" means that $A(B(x)) = x$ for each x.)

These two equalities show that the maps A and B are mutually inverse. Thus, the map A is invertible, i.e., bijective. In other words, A is a bijection. Hence, by the bijection principle,

$$|\mathcal{P}(S)| = \left|\{0,1\}^n\right| = |\{0,1\}|^n \qquad \text{(by (1))}$$
$$= 2^n,$$

so that Theorem 1.5.8 is proved again. □

A few exercises on the subject appear at the end of §1.5.3 in the 2019 notes.

1.6. Interchange of summations

1.6.1. The finite Fubini principle

Now let us return to the properties of finite sums. We have already seen some rules for manipulating such sums. Here is another:

Theorem 1.6.1 (finite Fubini principle)**.** Let X and Y be two finite sets. Let $a_{(x,y)}$ be a number for each pair $(x,y) \in X \times Y$. Then,

$$\sum_{x \in X} \sum_{y \in Y} a_{(x,y)} = \sum_{(x,y) \in X \times Y} a_{(x,y)} = \sum_{y \in Y} \sum_{x \in X} a_{(x,y)}.$$

The idea behind this theorem is that all three sides of this equality are just the sum of all the numbers $a_{(x,y)}$, organized in three different ways. This is easiest to see on an example in which $X = [n]$ and $Y = [m]$:

Example 1.6.2. Let $n, m \in \mathbb{N}$. Let $a_{(x,y)}$ be a number for each pair $(x,y) \in [n] \times [m]$. Then, the finite Fubini principle (applied to $X = [n]$ and $Y = [m]$) says that

$$\sum_{x \in [n]} \sum_{y \in [m]} a_{(x,y)} = \sum_{(x,y) \in [n] \times [m]} a_{(x,y)} = \sum_{y \in [m]} \sum_{x \in [n]} a_{(x,y)}.$$

In other words, it says that

$$\sum_{x=1}^{n} \sum_{y=1}^{m} a_{(x,y)} = \sum_{(x,y)\in[n]\times[m]} a_{(x,y)} = \sum_{y=1}^{m} \sum_{x=1}^{n} a_{(x,y)}. \tag{2}$$

This equality becomes very intuitive if you arrange the numbers $a_{(x,y)}$ into a rectangular table:

$$\begin{array}{cccc} a_{(1,1)} & a_{(1,2)} & \cdots & a_{(1,m)} \\ a_{(2,1)} & a_{(2,2)} & \cdots & a_{(2,m)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{(n,1)} & a_{(n,2)} & \cdots & a_{(n,m)} \end{array}.$$

Then, the equality (2) is simply saying that

$$\begin{aligned} &\sum_{x=1}^{n} (\text{sum of all entries of the } x\text{-th row of the table}) \\ &= (\text{sum of all entries of the table}) \\ &= \sum_{y=1}^{m} (\text{sum of all entries of the } y\text{-th column of the table}). \end{aligned}$$

This example shows that the finite Fubini principle is intuitively obvious (at least in the case when $X = [n]$ and $Y = [m]$; but the general case differs only in the way the numbers $a_{(x,y)}$ are labelled). A rigorous proof can be given by induction on $|X|$ (also, a reference appears in the 2019 notes).

The name "Fubini principle" is a reference to the (similar-looking, but much subtler) Fubini theorem for double integrals. Just as the latter makes it possible to interchange two integral signs, the former allows us to interchange two summation signs. When we use the Fubini principle to replace a sum of the form $\sum_{x=1}^{n} \sum_{y=1}^{m} a_{(x,y)}$ by $\sum_{y=1}^{m} \sum_{x=1}^{n} a_{(x,y)}$ (or vice versa), we say that we are **interchanging the summation signs**.

Here is a sample application of the Fubini principle (Exercise 1.6.1 in the 2019 notes):

Exercise 1. Let $n \in \mathbb{N}$. Let S be an n-element set. Find the sum of the sizes of all subsets of S, that is, find $\sum_{T\subseteq S} |T|$.

We shall solve this using the following simple (but extremely useful) fact:

Proposition 1.6.3 ("Counting by roll call")**.** Let S be a finite set.

(a) If T is a subset of S, then

$$|T| = \sum_{s \in S} [s \in T].$$

(b) For each $s \in S$, let $\mathcal{A}(s)$ be a logical statement (for example, "s is an even number" or "$s = \varnothing$"). Then,

$$(\text{\# of all } s \in S \text{ that satisfy } \mathcal{A}(s)) = \sum_{s \in S} [\mathcal{A}(s)].$$

Proof sketch. (See Proposition 1.6.3 in the 2019 notes for details.)

(a) The truth value $[s \in T]$ equals 1 for each $s \in T$ and equals 0 for each $s \in S \setminus T$. Hence, the sum $\sum_{s \in S} [s \in T]$ contains $|T|$ many addends equal to 1 (namely, one such addend for each $s \in T$) and $|S \setminus T|$ many addends equal to 0 (namely, one such addend for each $s \in S \setminus T$). Thus, this sum equals $|T| \cdot 1 + |S \setminus T| \cdot 0 = |T|$. This proves part **(a)**.

(b) Apply part **(a)** to the subset $T := \{s \in S \mid \mathcal{A}(s)\}$. Then,

$$|T| = (\text{\# of all } s \in S \text{ that satisfy } \mathcal{A}(s)),$$

and each $s \in S$ satisfies $[s \in T] = [\mathcal{A}(s)]$. Thus, part **(b)** follows from part **(a)**. $\square$

Example 1.6.4. Proposition 1.6.3 **(b)** (applied to $S = [5]$ and $\mathcal{A}(s) =$ ("s is odd")) yields

$$(\text{\# of } s \in [5] \text{ that are odd}) = \sum_{s \in [5]} [s \text{ is odd}] = 1 + 0 + 1 + 0 + 1.$$

Applied in the right situation, Proposition 1.6.3 can be very useful. Let us use it to solve Exercise 1:

Solution to Exercise 1. We have

$$\begin{aligned}
\sum_{T\subseteq S} \underbrace{|T|}_{\substack{=\sum_{s\in S}[s\in T] \\ \text{(by Proposition 1.6.3 } \textbf{(a)}\text{)}}} &= \sum_{T\subseteq S} \sum_{s\in S} [s\in T] \\
&= \sum_{s\in S} \underbrace{\sum_{T\subseteq S} [s\in T]}_{\substack{=(\text{\# of all subsets } T\subseteq S \text{ such that } s\in T) \\ \text{(by Proposition 1.6.3 } \textbf{(b)}\text{)}}} \qquad \left(\begin{array}{c}\text{here, we have interchanged the} \\ \text{summation signs, using the} \\ \text{finite Fubini principle}\end{array}\right) \\
&= \sum_{s\in S} \underbrace{(\text{\# of all subsets } T\subseteq S \text{ such that } s\in T)}_{\substack{=(\text{\# of all subsets of } S\setminus\{s\}) \\ \text{(since there is a bijection from } \{T\subseteq S \mid s\in T\} \\ \text{to } \{\text{subsets of } S\setminus\{s\}\} \text{ sending each } T \text{ to } T\setminus\{s\})}} \\
&= \sum_{s\in S} \underbrace{(\text{\# of all subsets of } S\setminus\{s\})}_{\substack{=2^{n-1} \\ \text{(by Theorem 1.5.8, since } S\setminus\{s\} \text{ is an } (n-1)\text{-element set)}}} \\
&= \sum_{s\in S} 2^{n-1} = \underbrace{|S|}_{=n} \cdot 2^{n-1} = n\cdot 2^{n-1}.
\end{aligned}$$

□

Remark 1.6.5. In the 2019 notes, I also give a different solution to Exercise 1, which proceeds by splitting the sum up according to the size $|T|$ of the subset T:

$$\sum_{T\subseteq S} |T| = \sum_{k=0}^{n} \underbrace{\sum_{\substack{T\subseteq S;\\ |T|=k}} |T|}_{\substack{=k\binom{n}{k} \\ \text{(why?)}}} = \sum_{k=0}^{n} k\binom{n}{k} = n\cdot 2^{n-1}$$

(by Proposition 1.3.33 in Lecture 8). Note that by confronting the two solutions against each other, we get a combinatorial proof of $\sum_{k=0}^{n} k\binom{n}{k} = n\cdot 2^{n-1}$.

1.6.2. The Fubini principle with a predicate

When applying the Fubini principle to rewrite a nested sum of the form $\sum_{x\in X}\sum_{y\in Y}$ into the form $\sum_{y\in Y}\sum_{x\in X}$, we say that we are **interchanging the summation signs**.

However, this works only if the ranges of the two sums are independent (i.e., the set Y does not depend on the choice of x). So we can interchange the summation signs in $\sum_{x=0}^{3} \sum_{y=0}^{5}$, but not in $\sum_{x=0}^{3} \sum_{y=0}^{x}$. To adapt the Fubini principle to cases like the latter, we need to add a simple tweak. Essentially, we extend the Fubini principle to "rectangular tables with gaps":

Theorem 1.6.6 (finite Fubini principle with a predicate)**.** Let X and Y be two finite sets. Let $\mathcal{A}(x,y)$ be a logical statement for each pair $(x,y) \in X \times Y$. Let $a_{(x,y)}$ be a number for each pair $(x,y) \in X \times Y$ that satisfies $\mathcal{A}(x,y)$. Then,

$$\sum_{x \in X} \sum_{\substack{y \in Y; \\ \mathcal{A}(x,y)}} a_{(x,y)} = \sum_{\substack{(x,y) \in X \times Y; \\ \mathcal{A}(x,y)}} a_{(x,y)} = \sum_{y \in Y} \sum_{\substack{x \in X; \\ \mathcal{A}(x,y)}} a_{(x,y)}.$$

For example, if $n, m \in \mathbb{N}$, then Theorem 1.6.6 (applied to $X = [n]$ and $Y = [m]$ and $\mathcal{A}(x,y) = (\text{"}x + y \text{ is even"})$) yields

$$\sum_{x \in [n]} \sum_{\substack{y \in [m]; \\ x+y \text{ is even}}} a_{(x,y)} = \sum_{\substack{(x,y) \in [n] \times [m]; \\ x+y \text{ is even}}} a_{(x,y)} = \sum_{y \in [m]} \sum_{\substack{x \in [n]; \\ x+y \text{ is even}}} a_{(x,y)}.$$

So we are allowed to interchange two summation signs even if the range of the inner one depends on the index of the outer one; we just have to make sure that the predicate that describes this dependence remains under the inner sum.

Here are some corollaries you can get by applying the principle to specific predicates:

Corollary 1.6.7 (triangular Fubini principle I)**.** Let $n \in \mathbb{N}$. For each pair $(x,y) \in [n] \times [n]$ with $x + y \leq n$, let $a_{(x,y)}$ be a number. Then,

$$\sum_{x=1}^{n} \sum_{y=1}^{n-x} a_{(x,y)} = \sum_{\substack{(x,y) \in [n] \times [n]; \\ x+y \leq n}} a_{(x,y)} = \sum_{y=1}^{n} \sum_{x=1}^{n-y} a_{(x,y)}.$$

Proof. The left hand side of this equality is

$$\sum_{x=1}^{n} \sum_{y=1}^{n-x} a_{(x,y)} = \sum_{x \in [n]} \sum_{\substack{y \in [n]; \\ y \leq n-x}} a_{(x,y)} = \sum_{x \in [n]} \sum_{\substack{y \in [n]; \\ x+y \leq n}} a_{(x,y)},$$

while the right hand side is

$$\sum_{y=1}^{n} \sum_{x=1}^{n-y} a_{(x,y)} = \sum_{y \in [n]} \sum_{\substack{x \in [n]; \\ x \leq n-y}} a_{(x,y)} = \sum_{y \in [n]} \sum_{\substack{x \in [n]; \\ x+y \leq n}} a_{(x,y)}.$$

Thus, the equality follows from Theorem 1.6.6 (applied to $X = [n]$ and $Y = [n]$ and $\mathcal{A}(x, y) = (\text{“}x + y \leq n\text{”})$). $\square$

Corollary 1.6.7 is called a "triangular Fubini principle" because it is about summing a triangular table of numbers:

$$\begin{array}{ccccc} a_{(1,1)} & a_{(1,2)} & \cdots & a_{(1,n-2)} & a_{(1,n-1)} \\ a_{(2,1)} & a_{(2,2)} & \cdots & a_{(2,n-2)} & \\ \vdots & \vdots & \iddots & & \\ a_{(n-2,1)} & a_{(n-2,2)} & & & \\ a_{(n-1,1)} & & & & \end{array}$$

(just like the original Fubini principle was about summing a rectangular table of numbers).

As an example for the use of the triangular Fubini principle, let us solve a simple exercise in a somewhat artistic way:

Exercise 2. Let $n \in \mathbb{N}$. Prove that

$$\sum_{k=1}^{n} k(n-k) = \binom{n+1}{3}.$$

Solution. We have

$$\sum_{k=1}^{n} k(n-k) = \sum_{x=1}^{n} x \underbrace{(n-x)}_{\substack{=\sum_{y=1}^{n-x} 1 \\ \text{(since a sum of } k \\ \text{many 1s equals } k)}} = \sum_{x=1}^{n} x \sum_{y=1}^{n-x} 1 = \sum_{x=1}^{n} \sum_{y=1}^{n-x} \underbrace{x \cdot 1}_{=x} = \sum_{x=1}^{n} \sum_{y=1}^{n-x} x$$

$$= \sum_{y=1}^{n} \underbrace{\sum_{x=1}^{n-y} x}_{\substack{=1+2+\cdots+(n-y) \\ =\frac{(n-y)(n-y+1)}{2} \\ =\binom{n-y+1}{2}}} \qquad \left(\begin{array}{c}\text{here, we interchanged the} \\ \text{summation signs using} \\ \text{Corollary 1.6.7}\end{array}\right)$$

$$= \sum_{y=1}^{n} \binom{n-y+1}{2} = \binom{n}{2} + \binom{n-1}{2} + \cdots + \binom{1}{2}$$

$$= \binom{1}{2} + \binom{2}{2} + \cdots + \binom{n}{2}$$

$$= \binom{0}{2} + \binom{1}{2} + \cdots + \binom{n}{2} \qquad \left(\text{since } \binom{0}{2} = 0\right)$$

$$= \binom{n+1}{2+1} \qquad \text{(by the hockey-stick identity)}$$

$$= \binom{n+1}{3},$$

qed. □