浅谈排列组合

前置知识

- 排列组合

$$A_n^m=\frac{n!}{(n-m)!}$$
$$\binom{n}{m}=C_n^m=\frac{A_n^m}{A_m^m}=\frac{n!}{(n-m)!\cdot m!}$$

特别的，当 $m>n$ 时，$C_n^m=0$ 。

- 排列组合的推导

在 n 个数里面 **按顺序** 的选择 m 个数，也就是 A_n^m ，我们可以这样假设：有 m 个空位，你需要一一把这 n 个物品放上去。有 n 种选择可以在第一个空位上面（ 因为每一个物品都可以放上去 ）。而第二个空位只有 $n-1$ 种选择，（因为有一个物品已经放在了第一个空位，剩下了 $n-1$ 个物品），第三个空位只有 $n-2$ 种选择 $\cdots\cdots$ ，第 m 个空位只有 $(n-m+1)$ 中选择。所以，$A_n^m=n\times(n-1)\times\cdots\times(n-m+1)=\dfrac{n!}{(n-m)!}$ 。

在 n 个数里面 **不按顺序** 的选择 m 个数，也就是 C_n^m ，也可以写成 $\binom{n}{m}$ 。如果在乎顺序选择，那么总结果就是 A_n^m 。如果不在乎顺序，就要去掉重复。所有被选出来的 m 个数，按顺序排列有 A_m^m 种，如果不按顺序排列，就只有一种。所以，$C_n^m=\dfrac{A_n^m}{A_m^m}=\dfrac{n!}{(n-m)!\cdot m!}$ 。

- 排列组合的扩充

由定义，我们可以发现 n 个数里面选取 m 个数的方案跟 n 个数里面选取 $n-m$ 个数的方案是一样的（因为 n 个数里面选 m 个数，就相当于在 n 个数里面选择 $n-m$ 个数丢掉），于是我们可以得到：

$$\binom{n}{m}=\binom{n}{n-m}$$

当在 $n+m$ 个数里面选择 m 个数（ $n\geq m$ ）的时候，相当于从 n 个数里面选择 $0,1,2,\cdots,m$ 个数，并且在剩下 m 个数里面选 $m,m-1,m-2,\cdots,0$ 个数的方案数之和。也就是说，

$$\sum_{i=0}^{m}\binom{n}{i}\binom{m}{m-i}=\binom{n+m}{m}$$

其中 $\sum$ 是求和符号，表示 $i=0,1,2,3,\cdots,m$ 时，求和符号右边式子的值的和。

当 $n=m$ 时，我们还可以得出

$$\sum_{i=0}^{n}\binom{n}{i}^2=\binom{2n}{n}$$

---

1. 组合数的递推式

在这一节，我们将会证明这个公式：

$$\binom{a}{b}=\binom{a-1}{b-1}+\binom{a-1}{b}$$

- 由 **定义** 出发

先来看看 $\binom{a}{b}$ 的含义。它指的就是从 a 个数里面不按顺序选出 b 个数的方案数。

我们可以这样分类讨论：

1. 不选最后一个数，那么就要从 $(a-1)$ 个数里面选出 b 个，即 $\binom{a-1}{b}$ 。
2. 选最后一个数，那么只要从 $(a-1)$ 个数里面选 $(b-1)$ 个数即可，那么就是 $\binom{a-1}{b-1}$ 。

由加法原理可以得到

$$\binom{a}{b}=\binom{a-1}{b-1}+\binom{a-1}{b}$$

- 由 **计算** 出发

$$\begin{aligned}
&\binom{a-1}{b-1}+\binom{a-1}{b}\\
=&\frac{(a-1)!}{(a-b)!\cdot(b-1)!}+\frac{(a-1)!}{(a-b-1)!\cdot b!}\\
=&\frac{(a-1)(a-2)\cdots(a-b+1)}{(b-1)!}+\frac{(a-1)(a-2)\cdots(a-b)}{b!}
\end{aligned}$$

提公因式 $(a-1)(a-2)\cdots(a-b+1)$，得原式

$$\begin{aligned}
&=(a-1)(a-2)\cdots(a-b+1)\cdot(\frac{1}{(b-1)!}+\frac{a-b}{b!})\\
&=(a-1)(a-2)\cdots(a-b+1)\cdot(\frac{b}{b!}+\frac{a-b}{b!})\\
&=(a-1)(a-2)\cdots(a-b+1)\cdot\frac{a}{b!}\\
&=\frac{a(a-1)(a-2)\cdots(a-b+1)}{b!}\\
&=\frac{a!}{(a-b)!\cdot b!}=\binom{a}{b}
\end{aligned}$$

$\square$

> 关于证明最后的小方块：QED 是拉丁词组 **Q**uod **E**rat **D**emonstrandum（这就是所要证明的）的缩写，代表证明完毕。现在的 QED 符号通常是 ■ 或者是 □。—— [OI-Wiki]

Ex.1 二项式定理

我们先从完全平方公式说起：

$$(a+b)^2=a^2+2ab+b^2$$

那么，$(a+b)^n$ 该怎么求呢？在这一节，我们将会证明这个公式，也就是 **二项式定理**：

$$\begin{aligned}
(a+b)^n&=\binom{n}{0}a^0b^n+\binom{n}{1}a^1b^{n-1}+\cdots+\binom{n}{n}a^nb^0\\
&=\sum_{i=0}^{n}\binom{n}{i}a^ib^{n-i}
\end{aligned}$$

其中 $\sum$ 是求和符号，表示 $i=0,1,2,3,\cdots,n$ 时，求和符号右边式子的值的和。

证明： 考虑 **数学归纳法**，

当 $n=0$ 时，$(a+b)^0=1\times a^0b^0=1$ 显然成立。

假设当 $n=k$ 时成立，即 $(a+b)^k=\sum_{i=0}^{k}\binom{k}{i}a^ib^{k-i}$ 。那么，当 $n=k+1$ 时，

$$\begin{aligned}
(a+b)^{k+1} &= (a+b)^k\times(a+b)\\
&= (a+b)\times(\binom{k}{0}b^k+\binom{k}{1}a^1b^{k-1}+\cdots+\binom{k}{k}a^k)\\
&= \binom{k}{k}a^{k+1}+\binom{k}{k-1}a^kb+\binom{k}{k-2}a^{k-1}b^2+\cdots+\binom{k}{0}ab^k\\
&\qquad\quad +\binom{k}{k}a^kb \quad +\binom{k}{k-1}a^{k-1}b^2+\binom{k}{k-2}a^{k-2}b^3+\cdots+\binom{k}{0}b^{k+1}\\
&= \binom{k}{k}a^{k+1}+(\binom{k}{k}+\binom{k}{k-1})a^kb+(\binom{k}{k-1}+\binom{k}{k-2})a^{k-1}b^2+\cdots+\binom{k}{0}b^{k+1}\\
&= \binom{k+1}{k+1}a^{k+1}+\binom{k+1}{k}a^kb+\binom{k+1}{k-1}a^{k-1}b^2+\cdots+\binom{k}{0}b^{k+1}\\
&= \sum_{i=0}^{k+1}\binom{k+1}{i}a^ib^{k+1-i}
\end{aligned}$$

$\square$

如果你不想看到这么复杂的计算过程，你也可以从它的组合意义出发，证明在这里。

当二项式定理中的 a 和 b 取特殊值时，你可以得到下面这些式子：

$$\begin{aligned}
\binom{n}{0}+\binom{n}{1}+\cdots+\binom{n}{n}=2^n &\qquad (a=b=1)\\
\sum_{i=0}^{n}(-1)^i\binom{n}{i}=0 &\qquad (a=-1,b=1)
\end{aligned}$$

---

2. 杨辉三角与排列组合

先将杨辉三角集体左移，放在一张表格上。（注意表格的行 / 列开头是 **0**）

行 / 列	0	1	2	3	4	5	6
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
6	1	6	15	20	15	6	1

我们设表格中第 x 行，第 y 列的数是 $P(x,y)$ 。

首先，由杨辉三角本身可以得到 $P(a,b)=P(a-1,b)+P(a-1,b-1)$ 。

接着，我们可以发现一个比较神奇的事实：$P(a,b)=\binom{a}{b}$ 。

证明：考虑 **数学归纳法**，

显然可得 $P(0,0)=\binom{0}{0}$ 。

发现杨辉三角的递推式 $P(a,b)=P(a-1,b)+P(a-1,b-1)$
与组合数的递推式 $\binom{a}{b}=\binom{a-1}{b-1}+\binom{a-1}{b}$ 形式是相同的，由此可以判断下列等式成立：

$$P(a,b)=\binom{a}{b}$$

我们再来回顾一下二项式定理：

$$(a+b)^n=\sum_{i=0}^{n}\binom{n}{i}a^ib^{n-i}$$

将它的系数从前往后列出来：$\binom{n}{0},\binom{n}{1},\binom{n}{2},\cdots,\binom{n}{n}$

由 $P(a,b)=\binom{a}{b}$ 可以得到：这些数字刚好就是杨辉三角第 n 行从前往后的数字！

而这，就是杨辉三角和组合数之间的关系。

Ex.2 Catalan 数

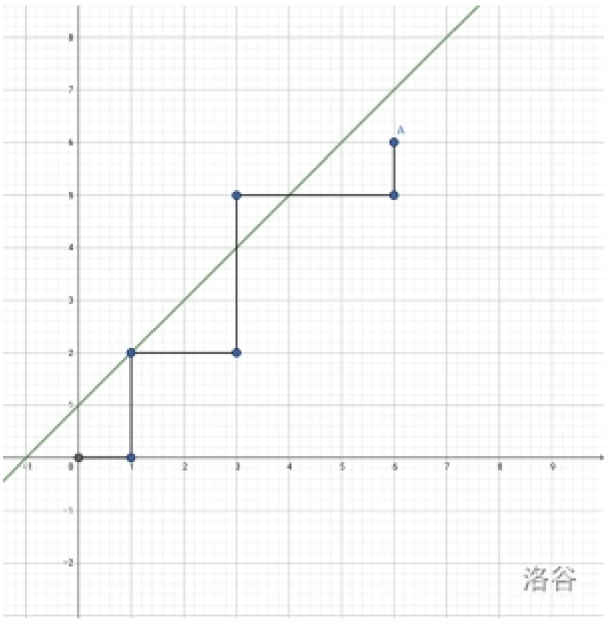
假如你在一个平面直角坐标系上，起点为 $(0,0)$ 。你一共可以移动 $2n$ 次，每次可以向右或向上走一格。我们规定，你在任意时刻往右走的次数都不能少于往上走的次数。问：当你走到点 (n,n) 时，有多少种不同的路径？

如果不按照规定来，方案则相当于在 $2n$ 次移动中选择 n 次向上移动，总数为 $\binom{2n}{n}$ 。

所有方案总数求出来了，我们接下来想一想不符合规定的方案总数。
先来想一个问题，什么时候方案是不符合要求的？

设你在移动若干次后的坐标为 (a,b) , 根据 "你在任意时刻往右走的次数都不能少于往上走的次数" 这一规定，可以得到这个不等式：$a\geq b$ 。也就是说，**我们在坐标系上，把 $y=x+1$ 这条线画出来，如果一条路径触碰到了这根线，它就不符合要求。**

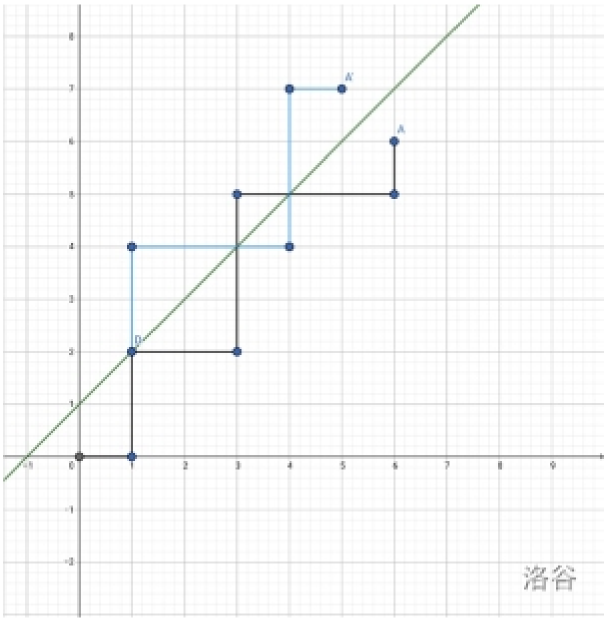
这种方案就是不符合要求的：



我们再尝试对这个不符合要求的方案转换一下：

所有的不符合要求的路径，都或多或少的触碰到了 $y=x+1$ 这一条直线。我们把第一次触碰到这条直线时的坐标记为 (a,b) 。

我们再把 (a,b) 之后的路径全都按照 $y=x+1$ 这条直线对称过去，此时图片变成了这样：
（黑色：原路径　蓝色：对称后路径）



对称过后，终点坐标就变成了 $(n-1,n+1)$ 。而且，所有不符合规定的路径对称后就唯一对应着一条到 $(n-1,n+1)$ 的路径。不仅如此，**所有到 $(n-1,n+1)$ 的一条路径也都唯一对应着一条不符合规定的路径**。（只需要再次对称回去就行了）从点 $(0,0)$ 到点 $(n-1,n+1)$ 的一条路径方案就相当于在 $2n$ 次移动操作中选取 $n+1$ 次的向上操作，总数为 $\binom{2n}{n+1}$ 。

那么，所有符合规定的路径方案总数为 $\dbinom{2n}{n}-\dbinom{2n}{n+1}$ 。

又因为 $\dbinom{a}{b}=\dbinom{a}{a-b}$ ，方案总数还可以表示为 $\dbinom{2n}{n}-\dbinom{2n}{n-1}$ 。

而这，就是 **Catalan 数** 的通项公式。这个问题只是 Catalan 数的一种形式。

Practice

1. 求证：$\sum_{i=0}^{n}\binom{n-i}{i}=F_{n+1}$ ，其中 F 指斐波那契数列。

证明：考虑 **数学归纳法**，

当 $n=0$ 时，$\dbinom{0}{0}=F_1=1$ 显然成立。当 $n=k$ 时，

$$
\begin{aligned}
F_k+F_{k-1}&=\sum_{i=0}^{k}\binom{k-i}{i}+\sum_{i=0}^{k-1}\binom{k-1-i}{i}\\
&=\binom{k}{0}+\binom{k-1}{1}+\binom{k-2}{2}+\binom{k-3}{3}+\cdots+\binom{1}{k-1}+\binom{0}{k}\\
&\qquad+\binom{k-1}{0}+\binom{k-2}{1}+\binom{k-3}{2}+\cdots+\binom{1}{k-2}+\binom{0}{k-1}\\
&=\binom{k+1}{0}+\binom{k}{1}+\binom{k-1}{2}+\cdots+\binom{2}{k-1}+\binom{1}{k}+\binom{0}{k+1}\\
&=\sum_{i=1}^{k+1}\binom{k+1-i}{i}=F_{k+1}
\end{aligned}
$$

□

2. Zapina / 有 n 个不同的人和 n 道不同的题，当第 i 个人开心时，当且仅当他被分配到 i 道题。请求出让至少一个人开心的方案数。

答案 = 总方案数 - 所有人都不开心的方案数

对于每道题而言，它都有 n 个选择来去到一个人的手中。而又有 n 道互不相同的题，根据乘法原理，所有总方案数自然就是 n^n 。我们设 $f(i,j)$ 为到了第 i 个人，分配了 j 道题目后都不开心的方案数，易得

$$f(i,j)=\sum_{k=0}^{j} f(i-1,k)\times\binom{j}{k}\times[k\neq i]$$

其中 $[k\neq i]$ 表示判断语句。当语句内容为真时，值为 1。如果为假，值为 0。
初始值 $f(0,0)=0$ ，终值为 $f(n,n)$ ，答案即为 $n^n-f(n,n)$ 。

Python 伪代码 / **C++ 关键代码**

3. **「雅礼集训2018」足球大战** / 有一场足球比赛，还有 n 秒就结束了，比分还是 $0:0$ 。主队每秒进球概率为 p ，客队每秒进球概率为 q ，求主队的获胜概率。

我们设 $T(i,j)$ 为队伍 i （ $i=1$ 表示主队，$i=0$ 表示客队 ）进球 j 次的概率。可以得到

$$P(\text{win})=\sum_{i=1}^{n}T(1,i)\sum_{j=0}^{i-1}T(0,j)$$

我们还可以得到，

$$T(i,j)=\begin{cases} p^j(1-p)^{n-j}\times\binom{n}{j} & (i=1)\\ q^j(1-q)^{n-j}\times\binom{n}{j} & (i=0)\end{cases}$$

我们又设 $S(k)=\sum_{i=0}^{k}T(0,i)$ ，可以得到如下递推式：

$$S(0)=T(0,0),\ S(n)=S(n-1)+T(0,n)\ (n\geq 1)$$

最后，再把这些式子整合在一起，可以得到

$$P(\text{win})=\sum_{i=1}^{n}T(1,i)\times S(i-1)$$

这里给出此题的 **C++ 正解代码** 。

All In All

我们学到了什么？

- 排列组合的定义，计算和拓展
- 二项式定理的公式和拓展
- 杨辉三角和排列组合的关系
- Catalan 数
- **数学归纳法**