组合数学知识点

排列组合

> 加法原理：完成一个工程有 n 种方法，其中 a_i 为第 i 种方法的数目，那么完成这件事有 $S=a_1+a_2+\cdots+a_n$ 种不同的方法。

> 乘法原理：完成一个工程有 n 个步骤，其中 a_i 为第 i 个步骤的不同方法数，那么完成这件事有 $S=a_1\times a_2\times\cdots\times a_n$ 种不同的方法。

组合数

从 n 个不同的元素中选择 $m(m\le n)$ 个元素构成的集合的方案数为 $\binom{n}{m}=\dfrac{n!}{m!(n-m)!}$。

二项式定理：$(a+b)^n=\sum\limits_{i=0}^{n}\binom{n}{i}a^ib^{n-i}$。

广义二项式定理：$(a+b)^\alpha=\sum\limits_{i=0}^{+\infty}\binom{\alpha}{i}a^{\alpha-i}b^i$，其中 $\binom{\alpha}{i}=\dfrac{\alpha^{\underline{i}}}{i!}$。

组合数的一些性质

1. $\binom{n}{m}=\binom{n}{n-m}$。
2. $\binom{n}{m}=\binom{n-1}{m-1}+\binom{n-1}{m}$（杨辉三角）。
3. $\sum\limits_{i=0}^{n}\binom{n}{i}=2^n$，$\sum\limits_{i=0}^{n}(-1)^i\binom{n}{i}=[n=0]$。
4. $\sum\limits_{i=1}^{k}\binom{n}{i}\binom{m}{k-i}=\binom{n+m}{k}$（范德蒙德恒等式）。
5. $\sum\limits_{i=0}^{n}i\binom{n}{i}=n2^{n-1}$。
6. $\sum\limits_{l=0}^{n}\binom{l}{k}=\binom{n+1}{k+1}$（朱世杰恒等式）。

二项式反演

如果 $g_n=\sum\limits_{i=0}^{n}\binom{n}{i}f_i$，那么就有 $f_n=\sum\limits_{i=0}^{n}(-1)^{n-i}\binom{n}{i}g_i$。

如果 $g_n=\sum\limits_{i=n}\binom{i}{n}f_i$，那么就有 $f_n=\sum\limits_{i=n}(-1)^{i-n}\binom{n}{i}g_i$。

容斥原理

若 $S=\{1,2,\dots,n\}$，有 n 个集合 $P_1,P_2\dots P_n\subseteq U$，那么有：

$\left|\bigcup\limits_{i=1}^{n}P_i\right|=\sum\limits_{T\subseteq S}(-1)^{|T|-1}\left|\bigcap\limits_{i\in S}P_i\right|$。

容斥原理一般化：

有两个对于集合的函数 $f(S)$ 和 $g(S)$：

如果 $g(S)=\sum_{T\subseteq S} f(T)$，那么 $f(S)=\sum_{T\subseteq S}(-1)^{|S|-|T|}g(T)$。

如果 $g(S)=\sum_{T\supseteq S} f(T)$，那么 $f(S)=\sum_{T\supseteq S}(-1)^{|T|-|S|}g(T)$。

发现二项式反演就是容斥原理的一种特殊形式。

min-max 容斥

对于满足**全序**关系并且其中元素满足可加减性的序列 $\{x_i\}$，设其长度为 n，并设 $S=\{1,2,3,\ldots,n\}$，则有：

$\max_{i\in S} x_i=\sum_{T\subseteq S}(-1)^{|T|-1}\min_{i\in T} x_i$。

$\min_{i\in S} x_i=\sum_{T\subseteq S}(-1)^{|T|-1}\max_{i\in T} x_i$。

这个式子常在求解期望的问题中使用。

kth min-max 容斥

记函数 kthmax 表示第 k 大的元素，同样定义 kthmin，则有：

$\operatorname{kthmax}_{i\in S} x_i=\sum_{T\subseteq S}(-1)^{|T|-k}\binom{|T|-1}{k-1}\min_{i\in T} x_i$。

$\operatorname{kthmin}_{i\in S} x_i=\sum_{T\subseteq S}(-1)^{|T|-k}\binom{|T|-1}{k-1}\max_{i\in T} x_i$。

斐波那契数列

斐波那契数列定义如下：$F_0=1,F_1=1,F_n=F_{n-1}+F_{n-2}(n\geq 2)$。

解析解：$F_n=\frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n-\left(\frac{1-\sqrt{5}}{2}\right)^n\right)$。

斐波那契数列的一些性质

1. $F_{n-1}F_{n+1}-F_n^2=(-1)^n$（卡西尼性质）。
2. $F_{n+k}=F_{n+1}F_k+F_nF_{k-1}$。若 $k=n$，则有 $F_{2n}=F_n(F_{n+1}+F_{n-1})$。
3. 有 2 性质归纳可得：$F_n\mid F_{nk}$。
4. 上述性质可逆，如果 $F_a|F_b$，则 $a|b$。
5. $\gcd(F_n,F_m)=F_{\gcd(n,m)}$。
6. 齐肯多夫定理：任何一个自然数 n 都可以唯一表示成一些斐波那契数的和：$n=F_{i_1}+F_{i_2}+\cdots+F_{i_k}$，其中 $i_j+2\leq i_{j+1},1\leq j<k$。

模意义下的周期性

考虑模 p 意义下的斐波那契数列，考虑这样 p^2+1 对斐波那契数：$(F_1,F_2),(F_2,F_3),\ldots,(F_{p^2+1},F_{p^2+2})$，根据抽屉原理，其中至少存在两对斐波那契数 $(F_a,F_{a+1}),(F_b,F_{b+1})$ 满足 $F_a=F_b,F_{a+1}=F_{b+1}$，这样的斐波那契数生成的下一个数是相同的，所以斐波那契数列是周期性的。

皮亚诺周期

模 m 意义下斐波那契数列的最小正周期被称为皮萨诺周期。皮萨诺周期总是不超过 $6m$，且只有在满足 $m=2\times 5^k$ 的形式时才取到等号。

错排

错位排列是没有任何元素出现在其有序位置的排列。即，对于 $1\sim n$ 的排列 P，如果满足 $P_i\neq i$，则称 P 是 n 的错位排列。记 D_n 为长度为 n 的错排的数量。

可以通过容斥得到：$D_n=n!\sum\limits_{i=0}^{n}\dfrac{(-1)^k}{k!}$。

也可以得到递推形式：$D_n=(n-1)(D_{n-1}+D_{n-2})$ 或 $D_n=nD_{n-1}+(-1)^n$。

甚至有一个很简单的取整形式：$D_n=\begin{cases}\left\lfloor\frac{n!}{e}\right\rfloor, & \text{if } n \text{ is odd}\\ \left\lceil\frac{n!}{e}\right\rceil, & \text{if } n \text{ is even}\end{cases}$。这也就说明了，错排的密度大约是 $\dfrac{1}{e}$ 的。

卡特兰数

考虑如下几个问题：

1. 长度为 $2n$ 的合法括号串的数量。
2. 一个无限大的栈，如果进栈序列为 $1,2,3,\ldots,n$，有多少种可能的出栈序列。
3. n 个节点有多少种不同的二叉树。

可以证明，这些问题的答案相同，均为 $H_n=\dfrac{\binom{2n}{n}}{n+1}$。

卡特兰数常见求值方式

$H_0=1,H_n=\sum\limits_{i=1}^{n}H_{i-1}H_{n-i},(n>0)$。

$H_n=H_{n-1}\times\dfrac{4n-2}{n+1}$。

$H_n=\dbinom{2n}{n}-\dbinom{2n}{n-1}$。

卡特兰数的封闭形式

卡特兰数的一大特点就是其生成函数的形式：假设其生成函数为 $H(x)$，可以通过递推式得知：

$H(x)=xH^2(x)+1$

解得 $H(x)=\dfrac{1\pm\sqrt{1-4x}}{2x}$，舍去不收敛的根之后，得到 $H(x)=\dfrac{1-\sqrt{1-4x}}{2x}$，将这个结果通过广义牛顿二项式展开之后就可以得到卡特兰数的通项公式。

斯特林数

第二类斯特林数

> 第二类斯特林数（斯特林子集数）：$\begin{Bmatrix} n \\ k \end{Bmatrix}$，也可记做 $S(n,k)$，表示将 n 个两两不同的元素，划分为 k 个互不区分的非空子集的方案数。

递推式：$\begin{Bmatrix} n \\ k \end{Bmatrix} = \begin{Bmatrix} n-1 \\ k-1 \end{Bmatrix} + k \begin{Bmatrix} n-1 \\ k \end{Bmatrix}$，$\begin{Bmatrix} n \\ 0 \end{Bmatrix} = [n=0]$。

通过容斥原理可以得到：$\begin{Bmatrix} n \\ k \end{Bmatrix} = \sum\limits_{i=1}^{k} \dfrac{(-1)^{k-i} i^n}{i!(k-i)!}$

第一类斯特林数

> 第一类斯特林数（斯特林轮换数）：$\begin{bmatrix} n \\ k \end{bmatrix}$，也可记做 $s(n,k)$，表示将 n 个两两不同的元素，划分为 k 个互不区分的非空轮换的方案数。

递推式：$\begin{bmatrix} n \\ k \end{bmatrix} = \begin{bmatrix} n-1 \\ k-1 \end{bmatrix} + (n-1) \begin{bmatrix} n-1 \\ k \end{bmatrix}$，$\begin{bmatrix} n \\ 0 \end{bmatrix} = [n=0]$。

第一类斯特林数没有实用的通项公式。

普通幂和上升幂之间的转化

记 $x^{\overline{n}} = \prod\limits_{i=0}^{n-1} (x+i)$。

那么就有：$x^{\overline{n}} = \prod\limits_{k=0}^{n} \begin{bmatrix} n \\ k \end{bmatrix} x^k$，$x^n = \prod\limits_{k=0}^{n} (-1)^{n-k} \begin{Bmatrix} n \\ k \end{Bmatrix} x^{\overline{k}}$。

普通幂和下降幂之间的转化

记 $x^{\underline{n}} = \prod\limits_{i=0}^{n-1} (x-i)$，注意到 $\dbinom{n}{m} = \dfrac{n^{\underline{m}}}{m!}$。

那么就有：$x^{\underline{n}} = \sum\limits_{k=0}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix} x^k$，$x^n = \prod\limits_{k=0}^{n} (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix} x^{\underline{k}}$。

贝尔数

记 B_n 表示将 n 个互不相同的元素划分成若干个不区分的集合的方案数。

递推式：$B_{n+1} = \sum\limits_{k=1}^{n} \dbinom{n}{k} B_k$。

同时 $B_n = \sum\limits_{k=0}^{n} \begin{Bmatrix} n \\ k \end{Bmatrix}$。

贝尔数本身没什么价值，但是 $B_{13} = 27644437, B_{14} = 190899322$，这也就意味着，和将 n 个元素划分成若干集合的问题中，暴力搜索的复杂度可以接受到 $n = 14$，一般可以在最终复杂度为 $O(3^n \operatorname{poly}(n))$ 或 $O(2^n \operatorname{poly}(n))$ 题目中获得可观的分数。

分拆数

> 分拆数 p_n：将 n 分成无序的若干个正整数的和的方案数。

> k 部分拆数 $p(n,k)$：将 n 分成无序的 k 个正整数的和的方案数。

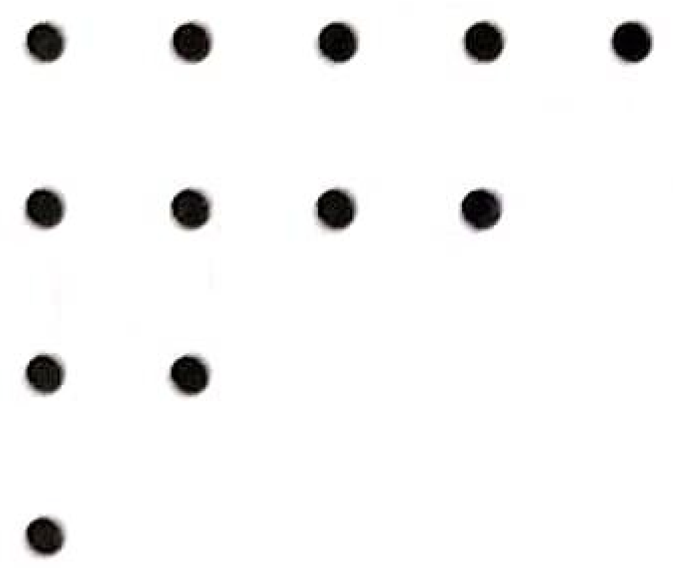
有递推式 $p(n,k)=\sum\limits_{j=0}^{k}p(n-k,j)$，$p(n,k)=p(n-1,k-1)+p(n-k,k)$。

Ferrers 图

Ferrers 图：将分拆的每个部分用点组成的行表示。每行点的个数为这个部分的大小。

根据分拆的定义，Ferrers 图中不同的行按照递减的次序排放。最长行在最上面。

例如：分拆 $12=5+4+2+1$ 的 Ferrers 图。



将一个 Ferrers 图沿着对角线翻转，得到的新 Ferrers 图称为原图的共轭，新分拆称为原分拆的共轭。显然，共轭是对称的关系。

例如上述分拆 $12=5+4+2+1$ 的共轭是分拆 $12=4+3+2+2+1$。

$O(n\sqrt{n})$ 求解分拆数

考虑求解 p_n，对于所有的拆分情况来说，选择 $>B$ 的正整数的个数不会超过 $\dfrac{n}{B}$ 个；其余的数都是 $\leq B$ 的数。

- 对于 $>B$ 的部分，先把每一个数扣掉 B，剩余的部分可以直接维护 $f_{i,j}$，表示和为 i，恰好 j 个数的分拆方案数。
- 对于 $\leq B$ 的部分，可以考虑其对应的共轭分拆，也可以呃直接维护 $g_{i,j}$ 表示和为 i，所有数都 $\leq j$ 的分拆方案数。

取 $B=\sqrt{n}$，两部分预处理出 f 和 g 的总时间复杂度最优，为 $O(n\sqrt{n})$。

此时，考虑 p_n 可以通过直接枚举 $>B$ 的数量 i 以及这一部分的和 S，那么 $p_n=\sum\limits_{i=1}^{\lfloor\frac{n}{B}\rfloor}\sum\limits_{S\leq n-iB}f_{S,i}g_{n-iB-S,B}$。

这样求解相较于直接使用生成函数或五边形数定理处理的优势在于，这样的处理可以在对拆分进行一定的限制之后，通过修改 f 和 g 的定义来完成处理。