

# 517 编程具体数学初级班之组合数学

吴一岐

2020 年 5 月 19 日

## 1 前言

组合数学是计算机科学中的重要组成部分，在信息学竞赛中经常与数论，概率，动态规划等需要计数的模型一起出现

并且，在各类复杂度分析，算法可行性分析，公式推导中，组合数学的内容也具有重要的支撑作用，因此，锻炼扎实的组合数学基本功是非常必要的！

# 目录

<b>1 前言</b>	<b>1</b>
<b>2 常见的一些符号表示的意义</b>	<b>4</b>
<b>3 基础知识</b>	<b>5</b>
3.1 组合数的阶乘表示	5
3.2 递推公式	5
3.3 二项式定理	5
3.4 二项式定理的变形	6
3.5 组合数公式常见变形	6
3.6 隔板法	7
3.7 错位排列	7
3.8 圆排列	7
3.9 多重集合的排列	7
3.10 多重集合的组合	8
3.11 鸽巢原理	8
3.11.1 普通鸽巢原理	8
3.11.2 加强版鸽巢原理	8
<b>4 康托展开</b>	<b>9</b>
4.1 计算排列的排名	9
4.2 计算特定排名的排列	9
4.3 下一个排列	9
<b>5 特殊的数列</b>	<b>10</b>
5.1 第一类斯特林数	10
5.2 第二类斯特林数	10
5.3 卡特兰数	11
5.3.1 卡特兰数的其他形式	12
5.3.2 卡特兰数的一些应用 (信息学中初赛经常用到)	12
<b>6 放球问题汇总</b>	<b>13</b>
6.1 球相同, 盒相同, 可以为空	13
6.2 球相同, 盒相同, 不能为空	13

6.3 球不同, 盒不同, 可以为空	14
6.4 球不同, 盒相同, 不能为空	14
6.5 球不同, 盒不同, 不能为空	14
6.6 球不同, 盒相同, 可以为空	14
6.7 球相同, 盒不同, 不能为空	14
6.8 球相同, 盒不同, 可以为空	14
<b>7 组合数取模</b>	<b>14</b>
7.1 第一阶段-普通阶乘逆元	15
7.2 第二阶段-卢卡斯定理	15
7.2.1 定理描述	15
7.2.2 定理证明	15
7.3 第三阶段-中国剩余定理合并	16

517编程

## 2 常见的一些符号表示的意义

1.  $a|b$  表示  $a$  是  $b$  的约数
2.  $\lfloor \frac{a}{b} \rfloor$ , 表示  $a$  除以  $b$  下取整
3.  $\prod_{i=1}^n p_i$  表示  $p_1, p_2, \dots, p_n$  的乘积
4.  $\sum_{i=1}^n a_i$  表示  $a_1, a_2, \dots, a_n$  的和
5.  $a \equiv b \pmod{p}$  表示  $a$  与  $b$  模  $p$  的结果是相同的
6.  $\sum_{d|n}$  表示枚举  $n$  的所有因子  $d$
7.  $f \times g(x)$  表示  $f$  函数与  $g$  函数的狄利克雷卷积
8.  $\int_a^b f(x)dx$  表示定积分
9.  $f(x) = [x = 1]$  表示  $f(x)$  只有在  $x = 1$  的时候才为 1, 其他情况为 0
10.  $\binom{n}{m}$  与  $C_n^m$  等价, 都表示  $n$  个不同的小球取  $m$  个方案数

## 3 基础知识

### 3.1 组合数的阶乘表示

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

### 3.2 递推公式

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

#### 预处理组合数

```
const int md = ; //模数
const int N = ; //数组大小
int c[N][N];
void init() {
    c[0][0] = 1;
    for (int i = 1; i < N; i++) {
        c[i][i] = c[i][0] = 1;
        for (int j = 1; j < i; j++) {
            c[i][j] = c[i-1][j] + c[i-1][j-1];
            c[i][j] %= md; // 一般N范围较大的题目都需要取模
        }
    }
}
```

### 3.3 二项式定理

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

### 3.4 二项式定理的变形

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

$$2. (x+y)^n = \sum_{k=0}^n \binom{n}{n-k} x^k y^{n-k}$$

$$3. (x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

$$4. \text{有一项为 } 1: (1+x)^n = \sum_{k=0}^n \binom{n}{k} x^k = \sum_{k=0}^n \binom{n}{n-k} x^k$$

$$5. \text{当 } x=2, y=1 \text{ 时即 } \binom{n}{0}2^0 + \binom{n}{1}2^1 + \dots + \binom{n}{n}2^n = 3^n$$

$$6. \text{当 } x=y=1 \text{ 时 } (x-y)^n = \binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \dots + (-1)^n \binom{n}{n} = 0 (n \geq 1)$$

$$7. \text{当 } x=y=1 \text{ 时 } (x+y)^n = \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n} = 2^n (n \geq 1)$$

$$8. \binom{n}{0} + \binom{n}{2} + \dots = 2^{n-1}$$

$$9. \binom{n}{1} + \binom{n}{3} + \dots = 2^{n-1}$$

### 3.5 组合数公式常见变形

$n, k$  都是正整数, 公式证明见视频

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

$$2. 1 \binom{n}{1} + 2 \binom{n}{2} + \dots + n \binom{n}{n} = n 2^{n-1}$$

$$3. n \binom{n-1}{0} + n \binom{n-1}{1} + \dots + n \binom{n-1}{n-1} = n 2^{n-1}$$

$$4. \sum_{k=1}^n k^2 \binom{n}{k} = n(n+1) 2^{n-2}$$

$$5. \sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}$$

### 3.6 隔板法

$n$  个相同的小球放到  $m$  个不同的盒子里, 盒子不能为空, 方案数为

$$\binom{n-1}{m-1}$$

上面这个也等价于  $x_1 + x_2 + \dots + x_m = n$  的正整数解的个数

$n$  个相同的小球放到  $m$  个不同的盒子里, 盒子可以为空, 可以先给每个盒子放一个球, 问题等价于将  $n+m$  个相同的小球放入  $m$  个不同的盒子里, 不为空

$$\binom{m+n-1}{m-1}$$

### 3.7 错位排列

每个数都不在自己位置的排列叫做错位排列, 假设排列为  $f_1, f_2 \dots f_n$ , 错位排列满足对于所有的  $1 \leq i \leq n, f_i \neq i$

错排的递推公式

$$\begin{aligned} f[1] &= 0, f[2] = 1 \\ f[i] &= (i-1) * (f[i-2] + f[i-1]) (i > 3) \end{aligned}$$

证明: 假设将  $i$  放到 1 到  $i-1$  的某个位置  $j$ , 假设  $j$  这个数放到  $i$  位置, 那么方案数就是剩下的  $i-2$  个数进行错排, 假如  $j$  这个数不放在  $i$  的位置上, 那么就是  $i-1$  个数进行错排, 注意这里思维的巧妙转换,  $i$  位置相当于  $j$  的一个新的限制位置了

### 3.8 圆排列

对于每种排列, 循环移动一圈, 都算同一种排列, 相当于将总的排列分成了若干组, 每组的本质是相同的, 而且每组里面有  $n$  个排列, 那么本质不同的圆排列就有  $\frac{n!}{n}$  种

### 3.9 多重集合的排列

设  $S$  是有  $k$  种不同类型对象的多重集合, 每一个元素都有无限重复数, 那么,  $S$  的  $r$  排列的数量是  $k^r$

设  $S$  是多重集合, 它有  $k$  种不同类型的对象, 且每一种类型的有限重复数分别为  $n_1, n_2, \dots, n_k$ . 设  $S$  的大小为  $n = n_1 + n_2 + \dots + n_k$ . 则  $S$  的排列数目等于

$$\frac{n!}{n_1! n_2! \dots n_k!}$$

证明: 相当于先选出  $n_1$  个位置, 再选出  $n_2$  个位置...

$$\binom{n}{n_1} \binom{n-n_1}{n_2} \binom{n-n_1-n_2}{n_3} \dots \binom{n-n_1-n_2-\dots-n_{k-1}}{n_k}$$

化简这个式子就可以证明了

### 3.10 多重集合的组合

设  $S$  是有  $n$  种类型的对象的多重集合, 每种元素均有无限的重复数, 那么  $S$  的  $k$  组合的个数等于

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

问题等价于选  $k$  次球, 每次都可以选  $n$  种球的方案数, 选球不分先后  
等价于  $k$  个相同的球放到  $n$  个不同的盒子里, 盒子可以为空

### 3.11 鸽巢原理

#### 3.11.1 普通鸽巢原理

如果要把  $n+1$  个物体放进  $n$  个盒子, 那么至少有一个盒子包含两个或更多的物体

比如, 在 13 个人中一定存在两个人生日是同一月份,  $n$  对新婚夫妇, 选择  $n+1$  人就一定会包含一对新婚夫妇。

#### 3.11.2 加强版鸽巢原理

设  $q_1, q_2, \dots, q_n$  是正整数, 如果将  $q_1 + q_2 + \dots + q_n - n + 1$  个物体放入  $n$  个盒子内, 那么或者第一个盒子至少含有  $q_1$  个物体, 或者第二个盒子至少含有  $q_2$  个物体, ..., 或者第  $n$  个盒子至少含有  $q_n$  个物体

证明: 采用反证法, 假设第  $i$  个盒子含有少于  $q_i$  个物体, 那么所有盒子中的物体总数不超过

$$(q_1 - 1) + (q_2 - 1) + \dots + (q_n - 1) = q_1 + q_2 + \dots + q_n - n$$



## 4 康托展开

### 4.1 计算排列的排名

给你一个排列, 计算这个排列之前的排列总数公式为

$$X = a_1(n-1)! + a_2(n-2)! + \dots + a_n 0!$$

$a_i$  是从左往右数的第  $i$  个数右边小于它的数的个数

比如  $[2, 3, 4, 1]$ , 那么  $a_1 = 1, a_2 = 1, a_3 = 1, a_4 = 0$

因此  $X = 1 * 3! + 1 * 2! + 1 * 1! + 0 = 9$ , 表示这个排列前面一共有 9 个排列, 这个排列是第 10 个排列

### 4.2 计算特定排名的排列

这个方法也经常叫康托展开的逆展开, 也就是告诉你一个排列的排名, 让你复原这个排列

比如让你求五个数的 107 个排列是谁, 假设第一个是 12345

假设排列的编号从 0 开始, 那么我们就是求编号为 106 的排列

利用康托展开的公式, 从高到低确定每一位

$\frac{106}{4!} = 4$ , 余 10, 说明第一位右边有 4 个比它小, 所以最高位为 5

$\frac{10}{3!} = 1$ , 余 4, 说明第二位右边有 1 个比它小, 所以当前位为 2

$\frac{4}{2!} = 2$ , 余 0, 说明第三位右边有 2 个比它小, 所以当前位为 4

$\frac{0}{1!} = 0$ , 余 0, 说明第四位右边没有数比它小, 所以当前位为 1

最后一位为 3

可以得到最终排列为  $[5, 2, 4, 1, 3]$

思想类似于进制转换

### 4.3 下一个排列

这里顺便提一下如何比较简单的寻找某个排列的下一个排列, 对于一个排列  $p_1, p_2, \dots, p_n$ , 我们从后往前扫描, 找到第一个  $p_i < p_{i+1}$  的地方, 然后从后往前找到第一个  $p_j < p_i$  将  $p_i$  与  $p_j$  交换, 然后将  $i$  位置以后的所有元素颠倒过来

比如  $[1, 2, 5, 7, 6, 4, 3]$ , 找到第一对  $p_i < p_{i+1}$  是  $(5, 7)$ , 从后往前找到一个数 6 比 5 要大, 交换后得到  $[1, 2, 6, 7, 5, 4, 3]$ , 然后将  $[7, 5, 4, 3]$  反转可得  $[1, 2, 6, 7, 3, 4, 5]$

## 5 特殊的数列

### 5.1 第一类斯特林数

将  $n$  个不同物体排成  $k$  个非空环的方案数

$$dp[n][k] = (n-1)dp[n-1][k] + dp[n-1][k-1]$$

要么自己成一个环, 要么凑到之前的某个环里, 凑到之前的环里一共有  $n-1$  种空位可以插入

#### 第一类斯特林数

```
void init(int n) {
    dp[0][0]=dp[1][1]=1;
    for(int i=2;i<=n;i++){
        for(int j=1;j<=i;j++){
            dp[i][j]=dp[i-1][j-1]+(i-1)*dp[i-1][j];
            dp[i][j]%=MOD;
        }
    }
}
```

### 5.2 第二类斯特林数

将  $n$  个不同物体分成  $k$  个非空集合的方案数

相当于将  $n$  个不同的小球, 放到  $k$  个相同的盒子里, 不能为空

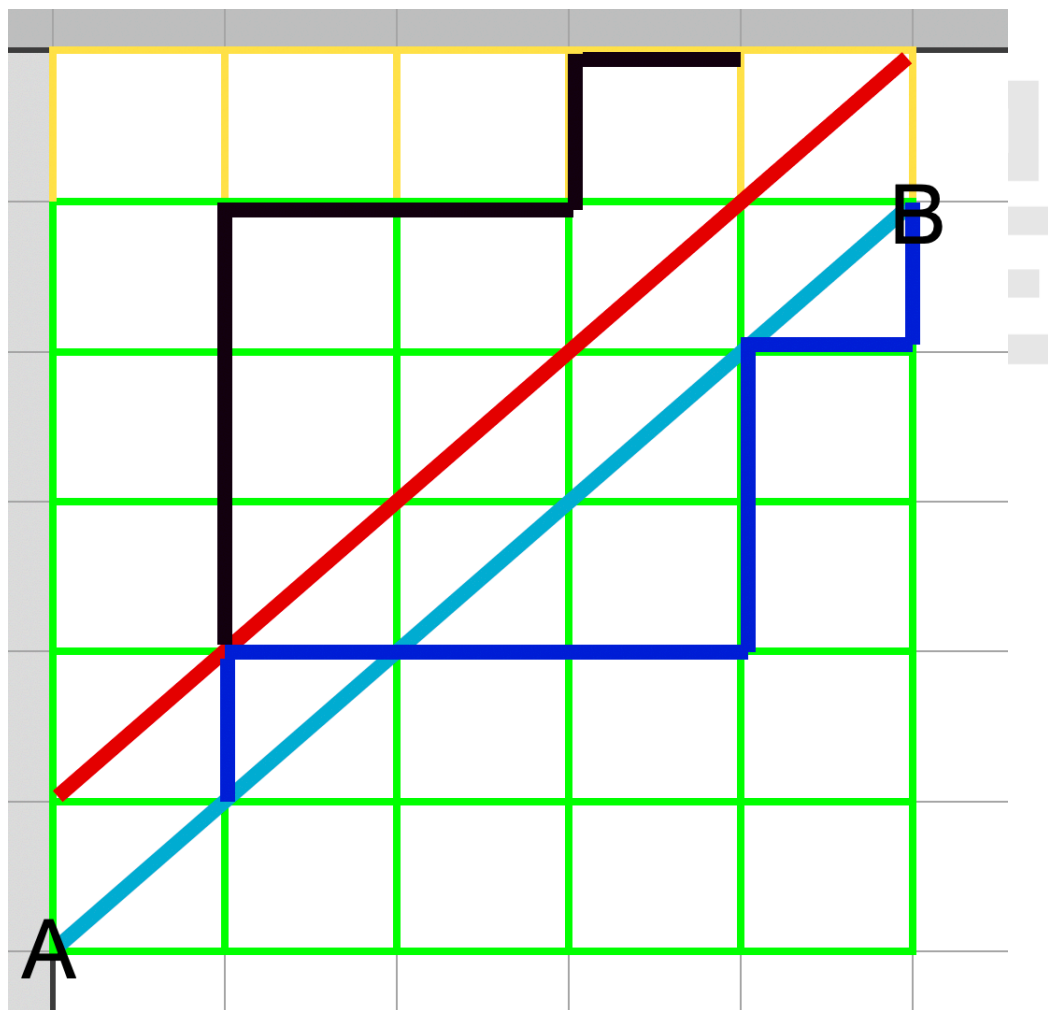
$$dp[n][k] = k * dp[n-1][k] + dp[n-1][k-1]$$

即, 当前这个球要么放到之前的盒子里, 要么新开一个盒子

## 第一类斯特林数

```
void init(int n){
    for(int i=0;i<=n;i++){
        dp[i][i]=dp[i][1]=1;
        for(int j=2;j<i;j++){
            dp[i][j]=(dp[i-1][j-1]+1LL*j*dp[i-1][j]%MOD)%MOD;
        }
    }
}
```

## 5.3 卡特兰数



假如有一个  $n * n$  的网格，从左下角的 A 点走到右上角的 B 点，始终不超过对角线的走法有多少种

观察上图可得，任何一条非法路径一定存在一个中间点在对角线上方，也就是会碰到红色的那条线，假设第一个碰到的是 P

我们从 P 点开始，将之后的那部分蓝色的线，按照对角线上方的斜线做对称，最终到达的点为  $(n-1, n+1)$

我们会发现任何一种非法方案都可以对称成  $(0, 0)$  到  $(n-1, n+1)$  的一条路径，他们之间存在一一映射关系，所以总方案数是  $\binom{2n}{n} - \binom{2n}{n+1}$

化简上面这个式子

$$\begin{aligned}
 Catalan(n) &= \binom{2n}{n} - \binom{2n}{n+1} \\
 &= \frac{(2n)!}{n!n!} - \frac{(2n)!}{(n+1)!(n-1)!} \\
 &= \frac{1}{n+1} \left( \frac{(2n)!(n+1)}{n!n!} - \frac{(2n)!}{n!(n-1)!} \right) \\
 &= \frac{1}{n+1} \left( \frac{(2n)!(n+1)}{n!n!} - \frac{(2n)!*n}{n!n!} \right) \\
 &= \frac{1}{n+1} \frac{(2n)!(n+1) - (2n)!*n}{n!n!} \\
 &= \frac{1}{n+1} \frac{(2n)!}{n!n!} \\
 &= \frac{1}{n+1} \binom{2n}{n}
 \end{aligned}$$

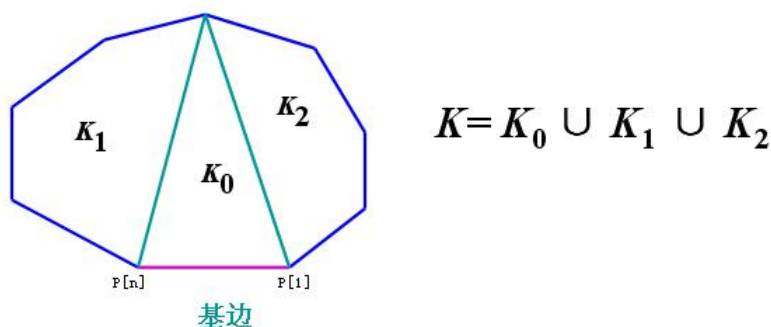
### 5.3.1 卡特兰数的其他形式

$$\begin{aligned}
 C(n) &= \frac{1}{n+1} \sum_{i=0}^n \binom{n}{i}^2 \\
 C(n+1) &= \sum_{i=0}^n C(i)C(n-i), C(0) = 1 \\
 C(n+1) &= \frac{2(2n+1)}{n+2} C(n), C(0) = 1
 \end{aligned}$$

### 5.3.2 卡特兰数的一些应用 (信息学中初赛经常用到)

1.  $n$  对括号的合法匹配的方案数
2.  $n$  个节点二叉树的形态数
3.  $n+1$  个叶子 ( $n$  个非叶节点) 的满二叉树 (每个节点都有两个儿子或者没有儿子) 的形态数, 走到左儿子  $+1$ , 走到右儿子  $-1$ , 类似于括号匹配. 2, 3 两种情况其实是一回事
4.  $n$  个数入栈后出栈的排列总数

5. 对凸  $n+2$  边形进行不同的三角形分割的方案数, 分割点不能交叉, 可以在端点处相交



选择一条基础边, 中间的三角形将整个多边形劈开两半, 变成了两个子问题, 就可以用卡特兰数的其他形式中的第二种形式来做了.

## 6 放球问题汇总

$n$  个球放到  $m$  个盒子里

### 6.1 球相同, 盒相同, 可以为空

等价于整数拆分问题, 比如将 7 拆分成小于等于 3 个数的方案有  $[1,1,5]$ ,  $[1,2,4]$ ,  $[1,6]$  等, 注意, 由于盒子是相同的, 所以  $[1,1,5]$  和  $[1,5,1]$  是同一种方案

我们可以用  $dp[i][j]$  表示  $i$  个球放到  $j$  个盒子里的方案数

转移方程为  $dp[i][j] = dp[i][j-1] + dp[i-j][j]$

两种转移分别表示

1: 放入小球后, 球最少的盒子为空, 相当于将  $i$  个小球放入  $j-1$  个盒子里,  $dp[i][j-1]$

2: 放入小球后, 球最少的盒子不为空, 说明每个盒子里面都有球, 可以从  $dp[i-j][j]$  转移过来

关于整数拆分的延伸阅读

<http://www.matrix67.com/blog/archives/6348>

### 6.2 球相同, 盒相同, 不能为空

不能为空就可以利用上面整数拆分的结果,  $dp[i][j] - dp[i][j-1]$

### 6.3 球不同, 盒不同, 可以为空

每个球都有  $m$  种不同的选择

答案为  $n^m$

### 6.4 球不同, 盒相同, 不能为空

$dp[i][j]$ , 将  $i$  个球放入  $j$  个盒子里的方案数

当前这个球一种选择是放入之前的某个盒子, 虽然盒子是相同的, 但是因为球是不同的导致当前这个球放入这  $j$  个盒子的某一个有  $j$  种选择

另一种选择是不放入之前的盒子, 放入一个新盒子

$$dp[i][j] = j * dp[i-1][j] + dp[i-1][j-1]$$

所以这个本质上就是第二类斯特林数, 等价于将  $i$  个不同的小球分成了  $j$  个非空集合的方案数

### 6.5 球不同, 盒不同, 不能为空

先假设盒子都是相同的, 问题就转换成了前一个问题, 最后因为盒子不同, 所以任意一种排列都是不同的。

答案等于  $dp[n][m] * m!$

### 6.6 球不同, 盒相同, 可以为空

还是跟第二类斯特林数一样, 可以为空就相当于枚举一下最终用了几个盒子求和即可

### 6.7 球相同, 盒不同, 不能为空

经典的隔板法,  $\binom{n-1}{m-1}$

### 6.8 球相同, 盒不同, 可以为空

$$\binom{m+n-1}{n-1}$$

## 7 组合数取模

求  $\binom{n}{m} \pmod{p}$

## 7.1 第一阶段-普通阶乘逆元

数据范围：  $n, m \leq 1e6$

$p$  是质数，直接利用阶乘逆元，费马小定理搞定

## 7.2 第二阶段-卢卡斯定理

数据范围：  $n, m \leq 1e18, p$  是质数  $p \leq 1e6$

### 7.2.1 定理描述

首先，我们将  $n, m$  写成  $p$  进制数的形式

$$n = (a_0 a_1 \dots a_k)_p$$

$$m = (b_0 b_1 \dots b_k)_p$$

那么

$$\binom{n}{m} \equiv \prod_{i=0}^k \binom{a_i}{b_i} \pmod{p}$$

也可以用递归的形式来表示这个式子

$$\binom{n}{m} \equiv \binom{n \% p}{m \% p} \binom{n/p}{m/p} \pmod{p}$$

我们发现如果上面的这个式子成立的话，我们就可以一直递归到  $n < p, m < p$  的时候再进行计算

### 7.2.2 定理证明

首先介绍一个引理：

$$(a + b)^p \equiv a^p + b^p \pmod{p}$$

证明：根据费马小定理  $(a + b)^p \equiv a + b \equiv a^p + b^p \pmod{p}$

因此

$$(1 + x)^p \equiv 1 + x^p \pmod{p}$$

假设  $n \% p = b$

$$(1 + x)^n \equiv (1 + x)^{\lfloor \frac{n}{p} \rfloor p} (1 + x)^b \pmod{p}$$

$$\equiv (1+x^p)^{\lfloor \frac{n}{p} \rfloor} (1+x)^b \pmod{p}$$

二项式展开后变成如下式子

$$\sum_{i=0}^k \binom{\lfloor \frac{n}{p} \rfloor}{i} x^{pi} \sum_{j=0}^k \binom{b}{j} x^j$$

注意这个式子与  $(1+x)^n$  的展开是等价的, 我们观察对于某一个  $x^m$  的系数, 在  $(1+x)^n$  里面是  $\binom{n}{m}$  而在上面这个式子中  $x^m$  被分成了两个部分, 假设  $m = i * p + j$ . ( $j < p$ ), 所以有

$$\binom{n}{m} = \binom{\lfloor \frac{n}{p} \rfloor}{i} \binom{b}{j} \quad (m = i * p + j, j < p)$$

$i = \lfloor \frac{m}{p} \rfloor, b = n \% p, j = m \% p$ , 所以有

$$\binom{n}{m} = \binom{\lfloor \frac{n}{p} \rfloor}{\lfloor \frac{m}{p} \rfloor} \binom{n \% p}{m \% p}$$

复杂度为  $O(\log_p^n)$

卢卡斯定理

---

```
int binom(int n, int m) {
    if (n < m) return 0;
    return 1LL * fac[n] * inv(fac[m]) % md * inv(fac[n - m]) % md;
}

int lucas(int n, int m) {
    if (m == 0) return 1;
    return 1LL * lucas(n / md, m / md) * binom(n % md, m % md) % md;
}
```

---

### 7.3 第三阶段-中国剩余定理合并

数据范围:  $n, m \leq 10^{18}, p \leq 10^6$  ( $p$  不一定是质数)

我们可以先将  $p$  进行质因数分解, 然后求组合数对于每个  $p_i^{a_i}$  取模的结果, 最后使用中国剩余定理进行答案的合并, 于是, 现在我们需要求



$$\binom{n}{m} \pmod{p_i^{a_i}}$$

将组合数写成阶乘的形式

$$\frac{n!}{m!(n-m)!}$$

由于分母可能包含质因子  $p_i$ , 于是不能直接求分母对  $p_i^{a_i}$  的逆元, 可以考虑将分子分母中包含的  $p_i$  全部提取出去, 最后再乘回去所以现在最关键的就是要求  $\frac{n!}{p_i^x} \pmod{p_i^{a_i}}$ , 即算出分子和分母的阶乘去除  $p_i$  因子后对  $p_i^{a_i}$  取模, 去掉之后就可以正常使用逆元

接下来以  $19! \pmod{3^2}$  为例

$$19! = 1 * 2 * 3 * \dots * 17 * 18 * 19$$

$$= 1 * 2 * 4 * 5 * 7 * 8 * 10 * 11 * 13 * 14 * 16 * 17 * 19 * (3 * 6 * 9 * 12 * 15 * 18)$$

$$= (1 * 2 * 4 * 5 * 7 * 8) * (10 * 11 * 13 * 14 * 16 * 17) * 19 * 3^6 * (1 * 2 * 3 * 4 * 5 * 6)$$

$3^6 * (1 * 2 * 3 * 4 * 5 * 6)$  相当于先提取出所有  $p_i$  的倍数, 然后, 每个数再提取出一个  $p_i$ , 最后变成  $n/p_i$  的阶乘, 这个可以递归求解

$(1 * 2 * 4 * 5 * 7 * 8)$  与  $(10 * 11 * 13 * 14 * 16 * 17)$  模  $p_i^{a_i}$  是相等的, 说明有循环节, 所以我们只需算出一个括号内的然后快速幂计算所有循环

最后剩下一部分零头的数量一定小于  $p_i^{a_i}$ , 可以暴力计算