二项式反演基础

清华大学 何昊天
kiana810@126.com

容斥原理

- 如何求若干个集合的并集的大小？
- 两个集合A,B：$|A\cup B|=|A|+|B|-|A\cap B|$
- 三个集合A,B,C：$|A\cup B\cup C|=|A|+|B|+|C|-|A\cap B|-|A\cap C|-|B\cap C|+|A\cap B\cap C|$
- n个集合：以此类推，可以写出一个统一的公式，粗糙地说，这个公式按照交集数量的奇偶性分配了正负号
- 注意到，为了计算并集的大小，我们只需要计算交集的大小就可以了，不严谨地说，并集和交集之间有某种对偶关系，这也是反演的核心所在：将具有某些对偶关系的函数相互转化

二项式定理

- 定理：$(a+b)^n=\sum_{i=0}^{n} C_n^i a^{n-i} b^i$
- 这个定理本身十分普通，也不再在这里讨论它的证明，我们考虑一种特殊情形，即a=1、b=-1的情况，再不妨定义0^0的值为1，我们将得到如下结果：

$$\sum_{i=0}^{n} (-1)^i C_n^i = e[n==0]$$

- 这个式子是接下来推导二项式反演的一个出发点，我们先结合容斥原理来理解一下其含义
- 考虑杨辉三角，假如我们依照容斥原理的奇偶性分配方式来给杨辉三角分配符号，即每行的第i(2|i)个数之前添一个负号，则除了第0行外，每一行的行和均为0，容斥原理通过这种巧妙的方式与二项式定理产生了联系

二项式反演引理

- 引理：$\sum_{i=k}^{n}(-1)^{n-i}C_n^iC_i^k=e[n==k]$
- 证明：$\sum_{i=k}^{n}(-1)^{n-i}C_n^iC_i^k=\sum_{i=k}^{n}(-1)^{n-i}C_n^kC_{n-k}^{i-k}$

$$=C_n^k\sum_{i=k}^{n}(-1)^{n-i}C_{n-k}^{i-k}=C_n^k*e[n==k]$$

- 这个引理的证明用到了刚才的结论，接下来我们就要利用这个引理来证明二项式反演定理

二项式反演定理

- 定理：$F(n)=\sum_{i=0}^{n}C_n^i f(i) \Leftrightarrow f(n)=\sum_{i=0}^{n}(-1)^{n-i}C_n^i F(i)$
- 证明：$\sum_{i=0}^{n}(-1)^{n-i}C_n^i F(i)=\sum_{i=0}^{n}(-1)^{n-i}C_n^i\sum_{k=0}^{i}C_i^k f(k)$

$$=\sum_{i=0}^{n}(\sum_{i=k}^{n}(-1)^{n-i}C_n^iC_i^k)f(i)=f(k)$$

- 这里也只给出从左到右的证明，从右到左的证明留给大家思考

二项式反演VS莫比乌斯反演

- 形式上来说，两个定理都是关于两个具有某种对偶性的函数的转化
- 不严格地说，莫比乌斯反演是更宽泛的，这也使得其相关应用更多
- 在竞赛中，二项式反演的情况大致如下：
  - 题目比莫比乌斯反演少，难度也小于莫比乌斯反演
  - 没有太多技巧，都是直接应用反演公式
  - 定位大概是较偏门的非主流算法，了解即可
- 接下来通过两道简单的例题帮助大家巩固二项式反演

错排问题

- 问题：试求有多少个长度为n的数列$a_1,a_2,\ldots,a_n$，满足$a_i \neq i$
- 这个题目非常经典，递推法或者公式法都可以解决，我们主要讨论如何用二项式反演得到错排问题的公式

错排问题 Solution

- 记长度为n的数列的错排数量为D_n
- 容易知道 $\sum_{i=0}^{n} C_n^i D_i = n!$
- 现在我们有较快的方法能够求出n!，但我们不知道D_n的值，又根据上式给出的关系，不难想到应用二项式反演：

$$D_n = \sum_{i=0}^{n} (-1)^{n-i} C_n^i * i!$$

$$= \sum_{i=0}^{n} (-1)^{n-i} \frac{n!}{(n-i!)} = \sum_{i=0}^{n} (-1)^i \frac{n!}{i!}$$

- 至此，错排问题可以O(n)解决了

染色问题

- 问题：有n个球排成一列，希望用k种颜色为它们染色，每种颜色至少要用依次，且要求相邻的球颜色不能相同，试求有多少种不同的染色方案
- 直接正面推公式也是可行的，但能够预感到会特别麻烦、特别复杂
- 任然考虑用二项式反演，不过需要一小步转化

染色问题 Solution

- 设当n固定时，原问题用k种颜色染色的答案为F_k
- 假设没有每种颜色至少用一次的限制，答案显然为$k(k-1)^{n-1}$
- 与上题类似，容易知道$k(k-1)^{n-1}=\sum_{i=0}^{k} C_k^i F_i$
- 同样的，$k(k-1)^{n-1}$很好算，但我们需要F_k的值，所以直接使用二项式反演：

$$F_k=\sum_{i=0}^{k}(-1)^{k-i}C_k^i * i(i-1)^{n-1}$$

- 至此，染色问题可以O(klogn)解决了

总结

- 额外介绍二项式反演，是因为二项式反演其实比莫比乌斯反演更简单一点，也是除了莫比乌斯反演外较常见的一种反演
- 如果想了解更多的反演，为大家推荐吕凯风同学的课件：
- http://vfleaking.blog.uoj.ac/slide/87#/

Thank you for listening!