莫比乌斯反演

——吉大附中实验学校 PoPoQQQ

什么是莫比乌斯反演？

- 介绍这个之前，我们先来看一个函数
- $$F(n)=\sum_{d|n} f(d)$$
- 根据F(n)的定义，我们显然有：
- F(1)=f(1)
- F(2)=f(1)+f(2)
- F(3)=f(1)+f(3)
- F(4)=f(1)+f(2)+f(4)
- F(5)=f(1)+f(5)
- F(6)=f(1)+f(2)+f(3)+f(6)
- F(7)=f(1)+f(7)
- F(8)=f(1)+f(2)+f(4)+f(8)

于是我们可以通过F(n)推导出f(n)

- f(1)=F(1)
- f(2)=F(2)-F(1)
- f(3)=F(3)-F(1)
- f(4)=F(4)-F(2)
- f(5)=F(5)-F(1)
- f(6)=F(6)-F(3)-F(2)+F(1)
- f(7)=F(7)-F(1)
- f(8)=F(8)-F(4)
- 在推导的过程中我们是否发现了一些规律?

莫比乌斯反演

- 公式：

$$F(n)=\sum_{d|n} f(d) \quad \Rightarrow \quad f(n)=\sum_{d|n}\mu(d)F(\frac{n}{d})$$

- 其中$\mu(d)$为莫比乌斯函数，定义如下：
- (1)若$d=1$则$\mu(d)=1$
- (2)若$d=p_1p_2...p_k$，p_i为互异素数，那么$\mu(d)=(-1)^k$
- (3)其它情况下$\mu(d)=0$

莫比乌斯函数的性质

- (1)对于任意正整数n有：

$$\sum_{d|n}\mu(d)=\begin{cases}1 & (n=1)\\ 0 & (n>1)\end{cases}$$

- 证明：
- ①当$n=1$时显然
- ②当$n\neq 1$时，将n分解可以得到$n=p_1^{a_1}p_2^{a_2}\dots p_k^{a_k}$
- 在n的所有因子中，μ值不为零的只有所有质因子次数都为1的因子，其中质因数个数为r个的因子有C_k^r个
- 那么显然有：

$$\sum_{d|n}\mu(d)=C_k^0-C_k^1+C_k^2+\dots+(-1)^k C_k^k=\sum_{i=0}^{k}(-1)^i C_k^i$$

莫比乌斯函数的性质

- 只需证明$\sum_{i=0}^{n}(-1)^i C_n^i = 0\ (n \in N_+)$即可
- 二项式定理：$(x+y)^n = \sum_{i=0}^{n} C_n^i x^i y^{n-i}$
- 令$x=-1, y=1$，代入即可得证。

莫比乌斯函数的性质

- (2)对于任意正整数n有：

$$\sum_{d|n}\frac{\mu(d)}{d}=\frac{\varphi(n)}{n}$$

- 其实没什么用，结论挺好玩的可以记一下
- 只需要令 $F(n)=n, f(n)=\varphi(n)$，代入莫比乌斯反演的公式即可
- 至于为什么 $n=\sum_{d|n}\varphi(d)$ 就留给大家做思考题了

莫比乌斯函数的性质

- (3)积性函数
- 数论上积性函数的定义：

  设$f(n)$为一个定义在$N+$集合上的函数，如果对于任意$(x,y)=1$有$f(xy)=f(x)f(y)$，则称$f(n)$为一个积性函数；若对于任意x和y均有$f(xy)=f(x)f(y)$，则称$f(n)$为一个完全积性函数
- 积性函数的性质：
- ① $f(1)=1$
- ②积性函数的前缀和也是积性函数

莫比乌斯函数的性质

- 由于莫比乌斯函数是积性函数，因此我们可以通过线性筛来求出莫比乌斯函数的值
- 代码：

```
mu[1]=1;
for(i=2;i<=n;i++)
{
        if(!not_prime[i])
        {
                prime[++tot]=i;
                mu[i]=-1;
        }
        for(j=1;prime[j]*i<=n;j++)
        {
                not_prime[prime[j]*i]=1;
                if(i%prime[j]==0)
                {
                        mu[prime[j]*i]=0;
                        break;
                }
                mu[prime[j]*i]=-mu[i];
        }
}
```

BZOJ 2440 完全平方数

- 题目大意：求第k个无平方因子数
- 无平方因子数(Square-Free Number)，即分解之后所有质因数的次数都为1的数
- 首先二分答案 问题转化为求[1,x]之间有多少个无平方因子数
- 根据容斥原理可知 对于sqrt(x)以内所有的质数 有
- x以内的无平方因子数
- =0个质数乘积的平方的倍数的数的数量(1的倍数)
- -每个质数的平方的倍数的数的数量(9的倍数,25的倍数,...)
- +每2个质数乘积的平方的倍数的数的数量(36的倍数,100的倍数,...)-...

BZOJ 2440 完全平方数

- 容易发现每个乘积a前面的符号恰好是$\mu(a)$ (例如$\mu(3)=-1$, 故9对答案的贡献为负；$\mu(6)=1$，故36对答案的贡献为正)
- x以内i^2的倍数有$\left\lfloor \frac{x}{i^2} \right\rfloor$个 故有 $Q(x)=\sum_{i=1}^{\lfloor \sqrt{x} \rfloor} \mu(i) \left\lfloor \frac{x}{i^2} \right\rfloor$
- 这题和莫比乌斯反演没关系，算是莫比乌斯函数的一个应用吧。。。

现在我们来证明莫比乌斯反演定理

$$F(n)=\sum_{d|n} f(d) \quad \Rightarrow \quad f(n)=\sum_{d|n}\mu(d)F(\frac{n}{d})$$

- 证明：

$$f(n)=\sum_{d|n}\mu(d)F(\frac{n}{d})=\sum_{d|n}\mu(d)\sum_{k|\frac{n}{d}} f(k)=\sum_{k|n} f(k)\sum_{d|\frac{n}{k}}\mu(d)= f(n)$$

- 这里利用到了 $\sum_{d|n}\mu(d)=\begin{cases}1 & (n=1)\\ 0 & (n>1)\end{cases}$ 这条性质

- 形式二：$F(n)=\sum_{n|d} f(d) \quad \Rightarrow \quad f(n)=\sum_{n|d}\mu(\frac{d}{n})F(d)$

- 证明同理 一般要用到的都是这种形式

有了这个定理，我们能干什么？

- 对于一些函数f(n)，如果我们很难直接求出它的值，而容易求出倍数和或约数和F(n)，那么我们可以通过莫比乌斯反演来求得f(n)的值
- 例：f(n)表示某一范围内(x,y)=n的数对的数量，F(n)表示某一范围内n|(x,y)的数对的数量
- 那么直接求f(n)并不是很好求，而F(n)求起来相对无脑一些，我们可以通过对F(n)进行莫比乌斯反演来求得f(n)
- 下面用几道例题来为大家讲解一下莫比乌斯反演的好处

BZOJ 2301 Problem b

- n次询问，每次询问有多少个数对(x,y)满足a<=x<=b,c<=y<=d且gcd(x,y)=k
- N<=5W,1<=a<=b<=5W,1<=c<=d<=5W
- 首先利用容斥原理将一个询问拆分成四个，每次询问有多少个数对(x,y)满足1<=x<=n,1<=y<=m且gcd(x,y)=k
- 这个问题等价于询问有多少个数对(x,y)满足1<=x<=floor(n/k),1<=y<=floor(m/k)且x与y互质
- 利用NOI2010能量采集中的方法，我们可以得到一个O(nlogn)的算法，但是这显然不能胜任此题的数据范围
- 这时候我们就可以考虑莫比乌斯反演了

BZOJ 2301 Problem b

- 由于之前的结论，我们可以令f(i)为1<=x<=n,1<=y<=m且gcd(x,y)=i的数对(x,y)的个数，F(i)为1<=x<=n,1<=y<=m且i|gcd(x,y)的数对(x,y)的个数
- 那么显然有 $F(i)=\left\lfloor \frac{n}{i} \right\rfloor \left\lfloor \frac{m}{i} \right\rfloor$
- 反演后可得 $f(i)=\sum_{i|d}\mu(\frac{d}{i})F(d)=\sum_{i|d}\mu(\frac{d}{i})\left\lfloor \frac{n}{d} \right\rfloor \left\lfloor \frac{m}{d} \right\rfloor$
- 枚举原题中k的每一个倍数，我们就可以O(n)时间处理每个询问了
- 但是O(n)还是不能胜任本题的数据范围
- 考虑进一步优化

BZOJ 2301 Problem b

- 观察式子，发现$\left\lfloor \frac{n}{d} \right\rfloor$最多有$2\sqrt{n}$个取值
- 那么$\left\lfloor \frac{n}{d} \right\rfloor \left\lfloor \frac{m}{d} \right\rfloor$就至多有$2(\sqrt{n}+\sqrt{m})$个取值
- 枚举这$2(\sqrt{n}+\sqrt{m})$个取值，对莫比乌斯函数维护一个前缀和，就可以在$\Theta(\sqrt{n})$时间内出解
- 总时间复杂度$\Theta(n\sqrt{n})$
- 枚举除法的取值这种方法在莫比乌斯反演的应用当中非常常用，且代码并不难写
- 不难写？怎么写？

BZOJ 2301 Problem b

- if(n>m) swap(n,m);
- for(i=1;i<=n;i=last+1)
- {
- last=min(n/(n/i),m/(m/i));
- re+=(n/i)*(m/i)*(sum[last]-sum[i-1]);
- }
- return re;
- 超级好写不是么?

BZOJ 2820 Y~~G~~Y的GCD

- 题目大意：求有多少数对(x,y)(1<=x<=n,1<=y<=m)满足gcd(x,y)为质数
- n,m<=1000W 数据组数<=1W
- 首先我们枚举每一个质数 那么答案就是

$$ans=\sum_{p}^{\min(n,m)}\sum_{d=1}^{\min(n,m)}\mu(d)\left\lfloor\frac{n}{pd}\right\rfloor\left\lfloor\frac{m}{pd}\right\rfloor$$

- 直接做显然TLE 考虑优化
- 令$pd=T$，那么有

$$ans=\sum_{T=1}^{\min(n,m)}\left\lfloor\frac{n}{T}\right\rfloor\left\lfloor\frac{m}{T}\right\rfloor\sum_{p|T}\mu(\frac{T}{p})$$

BZOJ 2820 Y~~G~~Y的GCD

- 如果能求出 $\sum_{p|T}\mu(\frac{T}{p})$ 的前缀和，这个问题就能在 $\Theta(\sqrt{n})$ 时间内出解。
- 只需要暴力枚举每一个质数，去更新这个质数的倍数即可。
- 由 $\lim_{n\to\infty}\sum_{i=1}^{n}\frac{1}{i}=\ln n+r$ 这个结论易知每个质数更新时是均摊

  $\Theta(\log n)$的，而质数个数恰好为$\Theta(n/\log n)$
- 故暴力枚举+维护前缀和的时间复杂度即为 $\Theta(n)$。

BZOJ 3529 数表

- 题目大意：令F(i)为i的约数和，q次给定n,m,a,求

$$\sum_{\substack{1<=i<=n\\1<=j<=m\\F(\gcd(i,j))<=a}} F(\gcd(i,j)) \bmod 2^{31}$$

- n,m<=10^5,q<=2W,a<=10^9
- a的限制十分讨厌 我们首先假设没有这个限制
- 令g(i)为1<=x<=n,1<=y<=m,gcd(x,y)=i的数对(x,y)的个数
- 那么显然有

$$g(i)=\sum_{i|d}\mu\left(\frac{d}{i}\right)\left\lfloor\frac{n}{d}\right\rfloor\left\lfloor\frac{m}{d}\right\rfloor$$

BZOJ 3529 数表

- F(i)利用线性筛可以在O(n)时间内处理出来 那么就有

$$ans=\sum_{i=1}^{\min(n,m)}F(i)g(i)=\sum_{i=1}^{\min(n,m)}F(i)\sum_{i|d}\mu(\frac{d}{i})\left\lfloor\frac{n}{d}\right\rfloor\left\lfloor\frac{m}{d}\right\rfloor=\sum_{d=1}^{\min(n,m)}\left\lfloor\frac{n}{d}\right\rfloor\left\lfloor\frac{m}{d}\right\rfloor\sum_{i|d}F(i)\mu(\frac{d}{i})$$

- 治好了我多年的公式恐惧症~~
- 现在我们只要有$\sum_{i|d}F(i)\mu(\frac{d}{i})$的前缀和就可以在$\Theta(\sqrt{n})$时间内解决这个弱化版的问题
- 与上一题相同，枚举每一个i，暴力更新i的倍数，然后处理前缀和，这样做是O(nlogn)的
- 那么现在有了a的限制怎么搞呢？

BZOJ 3529 数表

- 我们发现对答案有贡献的i只有F(i)<=a的i
- 我们离线处理，将询问按照a排序，i按照F(i)排序
- 每次询问将所有F(i)<=a的i按照之前的方式插入 用树状数组维护前缀和即可
- 时间复杂度$\Theta(n\log^2 n+q\sqrt{n}\log n)$
- 取模可以利用自然溢出int 最后再对2^31-1取与即可

BZOJ 2154 Crash的数字表格

- 题目大意：给定n,m,求 $\sum_{i=1}^{n}\sum_{j=1}^{m} lcm(i,j)$ (n,m<=10^7)

$$ans = \sum_{i=1}^{n}\sum_{j=1}^{m} lcm(i,j) = \sum_{i=1}^{n}\sum_{j=1}^{m} \frac{i \cdot j}{\gcd(i,j)}$$

- 枚举 $d = \gcd(i,j)$
- 令 $F(x,y) = \sum_{\substack{1<=i<=x \\ 1<=j<=y \\ \gcd(i,j)=1}} i \cdot j$
- 则有

$$ans = \sum_{d=1}^{\min(n,m)} \frac{d^2 \cdot F(\left\lfloor \frac{n}{d} \right\rfloor, \left\lfloor \frac{m}{d} \right\rfloor)}{d} = \sum_{d=1}^{\min(n,m)} d \cdot F(\left\lfloor \frac{n}{d} \right\rfloor, \left\lfloor \frac{m}{d} \right\rfloor)$$

BZOJ 2154 Crash的数字表格

- 继续令$Sum(x,y)=\sum_{i=1}^{x}\sum_{j=1}^{y}i\cdot j=\frac{x(x+1)}{2}\cdot\frac{y(y+1)}{2}$
- 那么根据莫比乌斯反演可以推出
$$F(x,y)=\sum_{i=1}^{\min(x,y)} i^2\cdot\mu(i)\cdot Sum(\left\lfloor\frac{x}{i}\right\rfloor,\left\lfloor\frac{y}{i}\right\rfloor)$$
- 不是很好推，和之前的思路一样，我不当堂推了
- 将两个式子分别进行 $\Theta(\sqrt{n})$的计算 可以得到一个$\Theta(\sqrt{n}\cdot\sqrt{n})=\Theta(n)$的算法
- 至此本题已经可以解决。

BZOJ 2693 jzptab

- 题目大意：同上题 多组数据
- 由于是多组数据 因此上一题的$\Theta(n)$算法显然超时
- 考虑进一步优化

$$ans=\sum_{d=1}^{\min(n,m)} d \sum_{i=1}^{\min(n,m)} i^2\cdot\mu(i)\cdot Sum(\left\lfloor\frac{n}{di}\right\rfloor,\left\lfloor\frac{m}{di}\right\rfloor)$$

$$=\sum_{D=1}^{\min(n,m)} Sum(\left\lfloor\frac{n}{D}\right\rfloor,\left\lfloor\frac{m}{D}\right\rfloor)\sum_{i|D}\frac{D}{i}\cdot i^2\cdot\mu(i)$$

- 观察后面的$\sum_{i|D}\frac{D}{i}\cdot i^2\cdot\mu(i)$，如果我们能对这个函数求出一个前缀和，那么就可以在$\Theta(\sqrt{n})$的时间内处理每个询问

BZOJ 2693 jzptab

- 注意到积性函数的约数和也是积性函数
- 因此后面的那坨东西可以利用线性筛求出来
- 线性筛当 $i \bmod prime_j \equiv 0$ 时不满足积性函数的条件，但是由于此时 $\mu(prime_j \cdot i) = 0$，故多出来的因数的函数值都是0，增加的只有原先因数的 $\frac{D}{i}$ 部分乘了个prime[j]而已
- 这两道题的公式都有些鬼畜，建议写这两道题之前先推推有关公式，权当治疗公式恐惧症了

谢谢大家