数论（前置知识预习课）

2024 年 11 月

feecle8146

前置知识目录

这里的内容都很简单。如果你对这份目录里的某个名词完全不了解，可以课前花一小段时间看看这个 PPT。

1. 求和号的定义
2. 唯一分解定理，试除法分解质因数
3. 简单的数论函数：约数 k 次幂和，欧拉函数，莫比乌斯函数，gcd 和 lcm 的定义和求法
4. 同余式的定义以及基本操作
5. 埃氏筛
6. exgcd 解方程 $ax + by = \gcd(x, y)$
7. 逆元的定义，费马小定理（欧拉定理）求逆元，exgcd 求逆元
8. 整除分块求 $\sum_{i \le n} floor\left(\frac{n}{i}\right)$
9. 阶的定义
10. 普通 BSGS（底数和模数互质）

认识求和号

求和号具有以下性质：

1. 可交换。$\sum_i \sum_j f(i,j) = \sum_j \sum_i f(i,j)$。可以理解成枚举一张数表，按行枚举和按列枚举是等价的。
   注意：交换求和号时，不要改变变量的取值范围。例如，
   $\sum_i \sum_{j \ge i} f(i,j) = \sum_j \sum_{i \le j} f(i,j)$。
2. 加法运算律：$\sum(a+b) = \sum a + \sum b$。
3. 乘法分配律：$\sum_i \sum_j f(i)g(j) = \sum_i f(i) \times \sum_j g(j)$。

在数论问题中，有时需要求出一个很长的带求和号的算式的值。此时，我们的目标一般来说是：

分离变量，使得各个变量间互不影响，从而用上面的2,3分解成子问题

分解质因数

每一个正整数都可以唯一写成

$$n = \prod_i p_i^{k_i}$$

的形式，其中 p_i 是严格递增排列的质数。该形式称为 n 的唯一分解。

可以使用试除法在 $O(\sqrt{n})$ 复杂度内求出 n 的质因数分解。

```
for (int i = 2; i * i <= n; i++) {
    if (n % i) continue;
    p[++cnt] = i;
    while (n % i == 0) n /= i, pk[cnt]++;
}
```

数论函数

约数个数

定义 $\sigma_0(n)$ 或 $d(n)$ 表示 n 的约数个数。可以用求和号表示为

$$\sigma_0(n) = \sum_{d|n} 1$$

若已知 m 和 n 的唯一分解分别为 $\prod p_i^{l_i}, \prod p_i^{k_i}$，则 $m|n$ 等价于 $\forall i, l_i \le k_i$。

请你推导用 k_i 表示 $\sigma_0(n)$ 的公式。

$$\sigma_0(n) = \prod_i (k_i + 1)$$

数论函数

定义 $\sigma_k(n)$ 表示 n 的约数的 k 次方和。可以用求和号表示为

$$\sigma_k(n) = \sum_{d|n} d^k$$

请你用等比数列求和公式推导用 p_i, k_i 表示 $\sigma_t(n)$ 的公式。

$$\sigma_1(n) = \prod_i \frac{1 - p_i^{k_i t+1}}{1 - p_i^t}$$

数论函数

gcd 和 lcm

定义 $\gcd(a,b)$ 表示 a,b 的最大公因数，也即 $\max\limits_{u|a,u|b} u$。
定义 $\operatorname{lcm}(a,b)$ 表示 a,b 的最小公倍数，也即 $\min\limits_{a|u,b|u} u$。

gcd 的唯一分解就是 a,b 的唯一分解中，把每个质数的幂次分别取 min。
lcm 的唯一分解就是 a,b 的唯一分解中，把每个质数的幂次分别取 max。

因为 $\min(a,b)+\max(a,b)=a+b$，所以 $\gcd(a,b)\times\operatorname{lcm}(a,b)=ab$。

数论函数

gcd 和 lcm

有结论：$u|a, u|b \to u|\gcd(a,b)$；$a|u, b|u \to \operatorname{lcm}(a,b)|u$。试着证明该结论。

同时，我们有 $\gcd(a,b) = \gcd(a, a+kb)\ (k \in \mathbb{Z})$。

因此，可以用辗转相除法求出两个数的 gcd，进而求出 lcm。

对于多个数的 gcd/lcm，只需按任意顺序依次求出。

```
int gcd(int a, int b) { return !b ? a : gcd(b, a % b); }
```

数论函数

gcd 和 lcm

注意到 $a|b \to a \leq b$，所以 $\gcd(a,b) \leq \min(a,b)$。进一步，若 $a \neq b$，则 $\gcd(a,b) \leq |a-b|$。

另一方面，若 $a \neq b$，则 $a|b \to a \leq \frac{b}{2}$，所以 $a \neq b \to \gcd(a,b) \leq \frac{\min(a,b)}{2}$。

这些简单的放缩，有时可能成为题目的突破口。

若 $\gcd(a,b) = 1$，就说 a, b 互质，有时记作 $a \perp b$。互质等价于唯一分解里没有共同因子。

$a|b$ 等价于 $\gcd(a,b) = a$。

数论函数

符号 $a \equiv b \pmod{c}$ 表示 $c|a-b$，也就是存在正整数 k 使得 $a = b + ck$。

若 $a \equiv b \pmod{c}, d|c$，则也有$a \equiv b \pmod{d}$。换句话说，模数更小的同余式能提供更多的信息。

同余意义下也可以执行加减乘操作。对于整除操作，若 $g|a, g|b$，则

$$\frac{a}{g} \equiv \frac{b}{g} \left(\bmod \frac{c}{\gcd(c, g)}\right)$$

数论函数

定义 $\varphi(n)$ 表示 $\leq n$ 的正整数中与 n 互质。可以用求和号表示为

$$\varphi(n) = \sum_{i \leq n} [i \perp n]$$

有下列计算 φ 的公式，它的本质是容斥原理：总的 – 钦定一个质因子的 + 钦定两个质因子的 - ...。

若 $n = \prod p_i^{k_i}$，则

$$\varphi(n) = n \times \prod_i (1 - \frac{1}{p_i}) = \prod_i (p_i^{k_i} - p_i^{k_i - 1})$$

数论函数

莫比乌斯函数

莫比乌斯函数 $\mu(n)$ 的定义较为绕口：

- 若 $\exists k_i \geq 2$，则 $\mu(n) = 0$。
- 否则，若 n 有 u 个质因子，则 $\mu(n) = (-1)^u$。

$\mu(n)$ 有如下性质：

$$\sum_{d|n} \mu(d) = [n = 1]$$

试着证明！
实际上就是二项式定理，或者也可以理解为 -1 和 1 相消。
莫比乌斯函数的重要用途是“莫比乌斯反演”，但我们今天不涉及。

筛法

埃式筛

我们希望求出 $2 \sim n$ 的每一个数是不是质数。

普通筛法的想法是，枚举每个正整数 i 的倍数 $2i, 3i, \ldots$ 并把他们标记为非质数。

思考：这段代码的时间复杂度如何表达？

可以写为 $\sum_i \frac{n}{i} = n \times \sum_{i \le n} \frac{1}{i} = O(n \log n)$ 。

不过，如果注意到只枚举质数作为 i 就够了，可以将复杂度变为 $O(n \log \log n)$，这个复杂度作为结论记住就够了。
此时的筛法被称为埃式筛。

```
for (int i = 2; i <= n; i++) {
    for (int j = i + i; j <= n; j += i) {
        vst[j] = 1;
    }
}
```

```
for (int i = 2; i <= n; i++) {
    if (vst[i]) continue;
    for (int j = i + i; j <= n; j += i) {
        vst[j] = 1;
    }
}
```

exgcd

exgcd

exgcd 的目的是解方程 $ax + by = \gcd(a, b)$。我们将构造性证明，该方程一定有解。

当 $b = 0$ 时，$(1,0)$ 是解；否则，假设 $bx' + (a \bmod b)y' = \gcd(a, b)$，令 $a \bmod b = a - cb$，左侧就是

$$x'b + y'a - cby' = ay' + b(x' - cy')$$

故只需递归求解 $(b, a \bmod b)$，再令 $x = y', y = x' - cy'$。

可以证明，$|x|, |y| \le \max(|a|, |b|)$。

```
void Exgcd(int a, int b, int &x, int &y) {
    if (!b) return x = 1, y = 0, void();
    int xx, yy;
    Exgcd(b, a % b, xx, yy);
    x = yy, y = xx - (a / b) * yy;
}
```

exgcd

逆元

exgcd 的目的是解方程 $ax+by=\gcd(a,b)$。

将 x 加上 $k\times b/\gcd(a,b)$，y 减去 $k\times a/\gcd(a,b)$，即可得到方程的全部解。由此可求出给定范围内的解数/最小解等，如 P5656。

特别地，当 $a\perp b$ 时，相当于 $ax\equiv 1 \pmod y$。此时，把 a 叫做 x 模 y 的逆元，写作 $a\equiv x^{-1} \pmod y$。

上述推导也说明，exgcd 是求逆元的一种方法。

exgcd

不定方程

exgcd 的目的是解方程 $ax+by=\gcd(a,b)$。

设 $\gcd(a,b)\mid c$，x,y 都乘上 $c/\gcd(a,b)$，就得到 $ax+by=c$ 的一组解；还可以用前述通解性质来缩小 x,y 的绝对值。

简单的例题：P1082，P2613

exgcd

逆元的性质

逆元具有唯一性。

逆元是一一对应关系，$(x^{-1})^{-1} = x$。

例子 (威尔逊定理)：$(p-1)! \bmod p$ 等于多少？

$p-1$。
除了 1 和 $p-1$，剩下的和逆元两两配对。

费马小定理和欧拉定理

费马小定理和欧拉定理

定理叙述如下：若底数和模数互质，则

$$x^{p-1} \equiv 1 \pmod p$$
$$x^{\varphi(n)} \equiv 1 \pmod n$$

这也说明在底数是质数时，x 的逆元就是 x^{p-2}。

我们还有所谓的“拓展欧拉定理”：任意 x，当 $b \geq \varphi(n)$ 时

$$x^b \equiv x^{(b \bmod \varphi(n))+\varphi(n)} \pmod n$$

换句话说，$\varphi(n)$ 是幂次取模后的值的循环节长度。（当然，不一定是最小循环节长度）

整除分块

整除分块的用途就是求 $\sum_{j\le n} floor\left(\frac{n}{j}\right)$ 这样的式子的值。为了简便，我们以后都省略 floor。事实上，它可以求任何 $\sum_{j\le n} g(j)\times f\left(\frac{n}{j}\right)$，其中 f 是好算的函数，而 $g(j)$ 的前缀和好算。

$\frac{n}{j}$ 可以有几种取值？提示：根号分治。

当 $j\le\sqrt{n}$ 时，j 只有 $\sqrt{n}$ 个；
当 $j>\sqrt{n}$ 时，$\frac{n}{j}\le\sqrt{n}$，只有 $\sqrt{n}$ 个。

因此，只有 $\le 2\sqrt{n}$ 种取值；换句话说，我们只需计算 $O(\sqrt{n})$ 种 $f(x)$，然后把每种 $f(x)$ 累加对应的次数。

整除分块

有两种实现方法：

1. 直接实现根号分治的代码，讨论 $x \le \sqrt{n}, x > \sqrt{n}$，并分别算出这个 x 对应几个 j。由于较为繁琐，在此不提。
2. 仍然枚举 j，但是注意到 $\frac{n}{j}$ 是分段相等且单调递减的，因此可以一次枚举一段。

我们（以及主流 OI 界）将采用第二种实现。这就需要知道，当我们枚举到 j 时，下一步应当跳到哪里，也就是是最大的 k 使得 $\frac{n}{k} = \frac{n}{j}$。

试着自己推出 k 的表达式？

$$\frac{n}{floor\left(\frac{n}{j}\right)}$$

整除分块

枚举 j，但是注意到 $\frac{n}{j}$ 是分段相等且单调递减的，因此可以一次枚举一段。

当我们枚举到 j 时，下一步应当跳到哪里，也就是最大的 k 使得 $\frac{n}{k}=\frac{n}{j}$。

$$k=\frac{n}{floor\left(\frac{n}{j}\right)}$$

```
for (int i = 1, j; i <= n; i = j + 1) {
    j = n / (n / i);
    ans += (j - i + 1) * f(n / i);
}
```

阶

阶

使得 $x^n \equiv 1 \pmod p$ 的最小正整数 n 称为 x 模 n 的阶，记作 $|x|$。

显然，阶存在的充要条件是 $\gcd(x, p) = 1$。

所有 $x^n \equiv 1 \pmod p$ 的 n 都是阶的倍数。这是因为阶也满足“辗转相除法”的性质。

结合拓展欧拉定理，可得 $|x| \mid \varphi(p)$。

阶是 x 的幂模 p 的最小循环节。由此可以推出，$\left|x^k\right| = \frac{|x|}{\gcd(k,|x|)}$。（可以画图理解）

BSGS

求解指数同余方程

指数同余方程的一般形式为，求下面方程 x 最小正整数解：

$$a^x \equiv b \pmod p$$

我们今天只解决 $\gcd(a,p)=1$ 的情况。

BSGS 算法的思想就是 meet in the middle。取块长 B，设 $x=iB-j$，
则 $a^x \equiv b$ 等价于 $(a^B)^i \equiv b \times a^j$。

将 $(a^B)^i$ 放入哈希表内（若有重复只保留 i 最小的），再依次查询 $b \times a^j$。

BSGS

求解指数同余方程

将 $(a^B)^i$ 放入哈希表内（若有重复只保留 i 最小的），再依次查询 $b \times a^j$。

由于 $x \le \varphi(p)$ 而 $\varphi(p) = O(p)$，所以取 $B = \sqrt{n}$ 可做到复杂度 $O(\sqrt{n})$。

特别地，如果固定 a, p 有多组 b，则只需要一次预处理（复杂度 $O\left(\frac{n}{B}\right)$），单次查询复杂度 $O(B)$，平衡复杂度后可做到总共 $O(\sqrt{nq})$。

由于在实数意义下，求 $a^x = b$ 的运算是“对数”，所以指数同余方程问题又称为离散对数问题。