数论

何俞均

gcd 相关

gcd 算是数论最基础的内容了。
不仅仅在数论题中，在其他题里面也有应用：

1. 以某个数为起点 gcd 不同的段数为 $O(\log n)$ 的。
2. 多个数求 gcd 时可以差分，这样子可以方便维护某些结构。

LOJ576 签到游戏

给出一个长度为 n 的序列 a，还有一个未知序列 b，你每次可以花费 $\gcd_{i=l}^{r} a_i$ 的代价得到 $\sum_{i=l}^{r} b_i$ 的值。
每次修改 a 中的一个数，求得到 b 中所有数字需要花费的最小权值。
$1 \leq n, q \leq 10^5, 1 \leq a_i \leq 10^9$

LOJ576 签到游戏 Solution

1. 首先把问题转化为生成树问题。
2. 考虑我们只有可能选择什么样的边。
3. 如何维护所需信息。

ICPC2022Hangzhou M.

一棵节点个数为 n 的无根带权树，给定一些关键点 $c_1, c_2 \dots c_k$，对于 n 个点中每一个点 i，求它这些关键点的距离和 s_i，以及其到这些关键点得到的 k 个距离数值的最大公约数 g_i。

$n \le 5 \times 10^5, k \le n, 1 \le w \le 10^7$

ICPC2022Hangzhou M. Solution

1. s_i 是好求的，只需考虑 $g_i = \gcd\{d_1, d_2, \cdots, d_k\}$ 怎么求。
2. 因为 $\gcd\{d_1, d_2, \cdots, d_k\} = \gcd\{d_1, d_2 - d_1, \cdots, d_k - d_1\}$，这样做的方便之处在于，我们每次一个点集的深度增加时，只需要改变 d_1 的大小。
3. 每棵子树维护二元组 (f, g) 分别表示第一项、后面的 gcd，从儿子到父亲的转移为 $(f + w, g)$，两棵子树合并就是 $(f_x, \gcd\{g_x, g_y, f_x - f_y\})$。
4. 把上面的 dp 换根就可以求 n 个点了。

CF1834E

给定长度为 n 的序列 a，求最小正整数 x，使得 x 不是 a 任何一个子段的 lcm。

$T \le 3 \times 10^5, \sum n \le 3 \times 10^5$

CF1834E Solution

1. 考虑答案可能出现的范围。
2. 类比 gcd，考虑 lcm 的出现特点是什么。
3. 如何维护这些 lcm 的值。

CCPC2023Harbin D

T 组询问，每次给定 $[l,r]$，求 $[l,r]$ 内的整数构成的最小生成树的大小，其中任意两数 (i,j) 之间的边权为 $\omega(lcm(i,j))$。$\omega(x)$ 为 x 的不同质因数个数。

$\sum r \leq 10^6$

CCPC2023Harbin D. Solution

1. 将所有的数按照质因数的无重集合分类，那么每个集合会优先向哪些集合连边。
2. 一种情况下是这些集合连完边之后剩下很多个质数单点，可以发现直接让它们互相连边就行了。
3. 另一种情况是区间内没有质数，这样子区间长度很小，直接 prim 即可。

整除分块

求 $\sum_{i=1}^{n} \left\lfloor \frac{n}{i} \right\rfloor$

Step 1：取值相同的为一段，只有 $\mathcal{O}(\sqrt{n})$ 个段。

证明：$i \le \sqrt{n}$ 时显然至多 $\sqrt{n}$ 种取值，$i > \sqrt{n}$ 时 $\left\lfloor \frac{n}{i} \right\rfloor < \sqrt{n}$，至多有 $\sqrt{n}$ 种取值。这样总共取值不超过 $2\sqrt{n}$ 个。

由于 $\left\lfloor \frac{n}{i} \right\rfloor$ 变化是连续的，所以有 $\mathcal{O}(\sqrt{n})$ 个段。

Step 2：求解

```
int ans = 0;
for (int i = 1, j; i <= n; i = j + 1) {
    j = n / (n / i);
    ans += (j - i + 1) * (n / i);
}
```

整除分块

上面代码的核心是找到最大的 j，满足 $\left\lfloor \frac{n}{i} \right\rfloor \le \left\lfloor \frac{n}{j} \right\rfloor$。简单推下：

$$\left\lfloor \frac{n}{i} \right\rfloor \le \left\lfloor \frac{n}{j} \right\rfloor$$

$$\left\lfloor \frac{n}{i} \right\rfloor \le \frac{n}{j}$$

$$j \le \frac{n}{\left\lfloor \frac{n}{i} \right\rfloor}$$

$$j = \left\lfloor \frac{n}{\left\lfloor \frac{n}{i} \right\rfloor} \right\rfloor$$

狄利克雷卷积

若 $\mathbf{h}=\mathbf{f}*\mathbf{g}$, 那么 $\mathbf{h}(x)=\sum_{d|x}\mathbf{f}(d)\mathbf{g}(\frac{n}{d})$
一些奇怪的函数关系式:

1. $\mathbf{d}=\mathbf{1}*\mathbf{1}$
2. $\sigma=\mathbf{1}*\mathbf{id}$
3. $\mathbf{id}=\mathbf{1}*\varphi$
4. $\sigma=\mathbf{1}*\mathbf{1}*\varphi=\mathbf{d}*\varphi$

狄利克雷卷积

若 $\mathbf{h} = \mathbf{f} * \mathbf{g}$, 那么 $\mathbf{h}(x) = \sum_{d|x} \mathbf{f}(d)\mathbf{g}(\frac{n}{d})$
一些奇怪的函数关系式:

1. $\mathbf{d} = \mathbf{1} * \mathbf{1}$
2. $\sigma = \mathbf{1} * \mathbf{id}$
3. $\mathbf{id} = \mathbf{1} * \varphi$
4. $\sigma = \mathbf{1} * \mathbf{1} * \varphi = \mathbf{d} * \varphi$

莫比乌斯反演

莫比乌斯函数 μ 的一个定义就是 $\mu * \mathbf{1} = \epsilon = [n = 1]$
展开就是 $\epsilon(n) = \sum_{d|n} \mu(d)$

然后还有就是 $\mu(n) = \begin{cases} 1 & n = 1 \\ 0 & n \text{ 含有平方因子} \\ (-1)^k & k \text{ 为} n \text{ 的本质不同质因子个数} \end{cases}$

莫比乌斯反演

莫比乌斯反演形式 1：

$$f(n) = \sum_{d|n} g(n) \Longrightarrow g(n) = \sum_{d|n} \mu(d) f\left(\frac{n}{d}\right)$$

简记为：

$$f = g * 1 \Leftrightarrow g = f * \mu$$

形式 2：

$$f(i) = \sum_{d=1}^{\lfloor \frac{n}{i} \rfloor} g(d \cdot i) \Longrightarrow g(i) = \sum_{d=1}^{\lfloor \frac{n}{i} \rfloor} \mu(d) f(d \cdot i)$$

入门题

1. 求 $\sum_{i=1}^{n}\sum_{j=1}^{m}[\gcd(i,j)=k]$
2. 求 $\sum_{i=1}^{n}\sum_{j=1}^{m}\gcd(i,j)^k$
3. 求 $\sum_{i=1}^{n}\sum_{j=1}^{m}d(ij)$

入门题 1

$$
\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{m}[\gcd(i,j)=k] &= \sum_{i=1}^{n/k}\sum_{j=1}^{m/k}[\gcd(i,j)=1] \\
&= \sum_{i=1}^{n/k}\sum_{j=1}^{m/k}\sum_{d|i,d|j}\mu(d) \\
&= \sum_{d=1}^{n/k}\mu(d)\left\lfloor\frac{n}{kd}\right\rfloor\left\lfloor\frac{m}{kd}\right\rfloor
\end{aligned}
$$

入门题 2

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{m}\gcd(i,j)^k &= \sum_{d=1}^{n} d^k \sum_{i=1}^{n}\sum_{j=1}^{m}[\gcd(i,j)=d] \\ &= \sum_{d=1}^{n} d^k \sum_{i=1}^{n/d}\sum_{j=1}^{m/d}[\gcd(i,j)=1] \\ &= \sum_{d=1}^{n} d^k \sum_{i=1}^{n/d}\sum_{j=1}^{m/d}\sum_{x\mid i,x\mid j}\mu(x) \\ &= \sum_{d=1}^{n} d^k \sum_{x=1}^{n/d}\mu(x)\left\lfloor\frac{n}{dx}\right\rfloor\left\lfloor\frac{m}{dx}\right\rfloor \\ &= \sum_{T=1}^{n}\left\lfloor\frac{n}{T}\right\rfloor\left\lfloor\frac{m}{T}\right\rfloor\sum_{d\mid T} d^k\mu\left(\frac{T}{d}\right)\end{aligned}$$

入门题 2

我们以这题为例探索线性筛积性函数的一般方法。首先求出 $g(p^c)$，这样根据质因数分解 $n=\prod p_i^{c_i}$ ，$g(n)=\prod g(p_i^{c_i})$，所有 g 都可以求出。
根据定义

$$
\begin{aligned}
g(p^c) &= \sum_{d|p^c} d^k \mu\left(\frac{p^c}{d}\right) \\
&= \sum_{i=0}^{c} p^{ik}\mu(p^{c-i}) \\
&= \sum_{i=c-1}^{c} p^{ik}\mu(p^{c-i}) \\
&= p^{kc}-p^{k(c-1)} = p^{k(c-1)}\left(p^k-1\right)
\end{aligned}
$$

入门题 2

所有的 $O(\frac{n}{\log n})$ 个质数都用快速幂求一下 g，其他的数 n 用最小素因子求 $f(n) = f(p^c)f(\frac{n}{p^c})$ 这样复杂度就是 $O(n + q\sqrt{n})$，q 是询问次数。

入门题 3

考虑将 $d(nm)$ 换一种表达方式：
$d(nm) = \sum_{x|n}\sum_{y|m}[\gcd(x,y)=1]$
带进去化式子可以得到：

$$\sum_{d=1}^{n}\mu(d)f\left(\left\lfloor\frac{n}{d}\right\rfloor\right)f\left(\left\lfloor\frac{m}{d}\right\rfloor\right), f(n)=\sum_{i=1}^{n}\left\lfloor\frac{n}{i}\right\rfloor$$

考虑 f 的意义：对于每个 i，答案加上 $[1,n]$ 中 i 的倍数个数。我们反向考虑约数就会发现 $f(n)=\sum_{i=1}^{n}d(i)$。

数论分块 $+$ 线性筛即可。复杂度 $O(n+\sqrt{n})$。

杜教筛

核心思想：求积性函数 f 前缀和 $S(n)$，找一个函数 g 使得 $f*g$ 和 g 的前缀和很好求。然后考虑一个式子：

$$\sum_{i=1}^{n}(f*g)(i)=\sum_{xy\le n}f(x)g(y)=\sum_{i=1}^{n}g(i)S\left(\left\lfloor\frac{n}{i}\right\rfloor\right)$$

通过移项变形得到：

$$S(n)=\frac{1}{g(1)}\left(\sum_{i=1}^{n}(f*g)(i)-\sum_{i=2}^{n}g(i)S\left(\left\lfloor\frac{n}{i}\right\rfloor\right)\right)$$

如果我们用上整除分块的技巧递归求并记忆化，复杂度是多少呢？

杜教筛

首先递归求 $S(k)$ 时 k 一定是个 $k \in [1, \lfloor\sqrt{n}\rfloor]$ 或 $k = \lfloor\frac{n}{i}\rfloor, i < \sqrt{n}$ 的形式。递归只展开一层，总复杂度为：
$O\left(\sum_{i=1}^{\sqrt{n}} \sqrt{\left\lfloor\frac{n}{i}\right\rfloor} + \sum_{i=1}^{n} \sqrt{i}\right) \approx O\left(\int_1^{\sqrt{n}} \frac{\sqrt{n}}{\sqrt{x}} \mathrm{d}x + \int_1^{\sqrt{n}} \sqrt{x} \mathrm{d}x\right) = O\left(n^{0.75}\right)$ 如果我们线性筛预处理了 $S(1), S(2), \ldots, S(n^c)$，其中 $c > \frac{1}{2}$，那么复杂度又是多少呢？

杜教筛

首先 $S(1), S(2), ..., S(\sqrt{n})$ 不用求了，其次求 $S\left(\left\lfloor\frac{n}{1}\right\rfloor\right), S\left(\left\lfloor\frac{n}{2}\right\rfloor\right), ..., S\left(\left\lfloor\frac{n}{\sqrt{n}}\right\rfloor\right)$ 会变成只需求 $S\left(\left\lfloor\frac{n}{1}\right\rfloor\right), S\left(\left\lfloor\frac{n}{2}\right\rfloor\right), ..., S\left(\left\lfloor\frac{n}{n^{1-c}}\right\rfloor\right)$。
那么复杂度为:

$$O\left(n^c + \sum_{i=1}^{n^{1-c}} \sqrt{\left\lfloor\frac{n}{i}\right\rfloor}\right) \approx O\left(n^c + \int_1^{n^{1-c}} \frac{\sqrt{n}}{\sqrt{x}} \mathrm{d}x\right) = \mathcal{O}\left(n^c + n^{1-0.5c}\right)$$

根据小学数学简单知识，$\max(c, 1-0.5c)$ 是一个增函数一个减函数取 max，在二者相等时该 max 最小。即 $c = \frac{2}{3}$ 时复杂度最小，为 $O(n^{\frac{2}{3}})$。

[模板] 杜教筛

求

$$S_1(n) = \sum_{i=1}^{n} \varphi(n)$$

$$S_2(n) = \sum_{i=1}^{n} \mu(n)$$

$n \leq 2^{31}$

[模板] 杜教筛

对 φ，选择 $\varphi * \mathbf{1} = \mathbf{id}$
对 μ，选择 $\mu * \mathbf{1} = \epsilon$
套用上面的模板就可以了。

洛谷 P3768 简单的数学题

求

$$\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\gcd(i,j)ij\right) \bmod p$$

$n \le 10^{10}, p \le 1.1\times 10^9, p$ 为质数

洛谷 P3768 简单的数学题 Solution

$$\begin{aligned}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\gcd(i,j)ij\right) &= \sum_{d=1}^{n} d^3 \sum_{i=1}^{\lfloor\frac{n}{d}\rfloor}\sum_{j=1}^{\lfloor\frac{n}{d}\rfloor}[\gcd(i,j)=1]ij \\ &= \ldots \\ &= \sum_{T=1}^{n}\left(1+2+\cdots+\lfloor\frac{n}{T}\rfloor\right)^2 T^2\varphi(T)\end{aligned}$$

令 $f(n)=n^2\varphi(n)$，$g(n)=n^2$ 即可。

[NOI2016] 循环之美

求

$$\sum_{i=1}^{n}\sum_{j=1}^{m}[i \perp j][j \perp k]$$

$1 \leq n, m \leq 10^9, 1 \leq k \leq 2000$

[NOI2016] 循环之美 Solution

$$
\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{m}[i\perp j][j\perp k] &= \sum_{j=1}^{m}[j\perp k]\sum_{i=1}^{n}\sum_{d|i,d|j}\mu(d) \\
&= \sum_{d=1}^{n}\mu(d)\lfloor\frac{n}{d}\rfloor\sum_{j=1}^{\lfloor m/d\rfloor}[jd\perp k] \\
&= \sum_{d=1}^{n}\mu(d)\lfloor\frac{n}{d}\rfloor\sum_{j=1}^{\lfloor m/d\rfloor}[j\perp k][d\perp k] \\
&= \sum_{d=1}^{n}\mu(d)\lfloor\frac{n}{d}\rfloor[d\perp k]\sum_{j=1}^{\lfloor m/d\rfloor}[j\perp k] \\
&= \sum_{d=1}^{n}\mu(d)\lfloor\frac{n}{d}\rfloor[d\perp k]F(\lfloor\frac{m}{d}\rfloor)
\end{aligned}
$$

[NOI2016] 循环之美 Solution

$F(n)$ 可以直接用 φ 函数求出，$F(n) = \frac{n}{k}\varphi(k) + F(n \bmod k)$。
设 $g(n,k) = \sum_{d=1}^{n}[d \perp k]\mu(d)$，那么

$$
\begin{aligned}
g(n,k) &= \sum_{d=1}^{n}\mu(d)\sum_{e|d,e|k}\mu(e) = \sum_{e|k}\mu(e)\sum_{i=1}^{\lfloor n/e\rfloor}\mu(ie) \\
&= \sum_{e|k}\mu(e)\sum_{i=1}^{\lfloor n/e\rfloor}[i \perp e]\mu(i)\mu(e) = \sum_{e|k}\mu(e)^2\sum_{i=1}^{\lfloor n/e\rfloor}[i \perp e]\mu(i) \\
&= \sum_{e|k}\mu(e)^2 g(\lfloor\frac{n}{e}\rfloor, e)
\end{aligned}
$$

边界情况为 $g(0,k) = 0$，$g(n,1)$ 需要杜教筛 μ。

Min_25 筛

1. Min_25 筛是解决一类积性函数求和问题，该积性函数满足 $f(p)$ 是一个关于 p 的低次多项式，且 $f(p^k)$ 可以快速计算。
2. 时间复杂度有说是 $O(n^{1-\epsilon})$，也有说是 $O\left(\dfrac{n^{0.75}}{\log n}\right)$，我也不会证明，我们就暂且认为是后者。
3. 个人认为 Min_25 筛的思路是：把 f 在 p^k 的点值拆成若干完全积性函数的加权和，进而处理出质数 f 之和，最后通过提出合数的最小素因子 p^e 得出所有函数值之和。

Min_25 筛

约定几个符号：$\mathrm{prime}(n)$ 表示 $\le n$ 的质数集合。$\mathrm{minp}(x)$ 表示 x 的最小质因子。p_i 表示第 i 个质数，特别地 $p_0 = 0$。符号 $\vee$ 表示或。

Step 1：分质数、合数与 1 两部分

$$\sum_{p\in\mathrm{prime(n)},p=2}^{n} f(p) + \sum_{i\notin\mathrm{prime(n)},i=1}^{n} f(i)$$

Step 2：预处理 g 数组

定义函数 $g_k(n,j)$ 表示 n 以内所有质数或最小素因子 $> p_j$ 的数的 k 次方和，我们后面将 k 省略不写。预处理这个主要是我们要把多项式 $f(x)$ 拆开一项一项求，那么

$g(n,j) = \sum_{i=2}^{n}[i \in \mathrm{prime(n)} \vee \mathrm{minp}(i) > p_j]i^k$

Min_25 筛

考虑如 DP 一样求 g。从 $g(n,j-1)$ 转移到 $g(n,j)$，数字显然是下降了。具体来说，所有满足 $\mathrm{minp}(i)=p_j$ 的合数 i，答案减去了 i^k 。即：

$$g(n,j)=g(n,j-1)-\begin{cases}0, p_j^2>n \\ p_j^k\left(g\left(\left\lfloor\frac{n}{p_j}\right\rfloor, j-1\right)-g\left(p_{j-1}, j-1\right)\right), p_j^2\le n\end{cases}$$

解释一下这个式子，转移时，$p_j^2>n$ 显然不存在合数最小素因子是 p_j。否则提出一个 i^k，剩下的数要求最小质因子 $\ge p_j$，当然还要减去 p_1 到 p_{j-1} 的情况，就是 $g(p_{j-1},j-1)$。当然在实际写代码时先处理好所有 $i\le\sqrt{n}$ 的 $\mathrm{sp}(i)=\sum_{j\le i}p_j$。

Min_25 筛

实际实现时，我们只计算到最大的 j 满足 $p_j^2 \le n$，因为 $g(n, \ge j)$ 的值相同。
巧妙之处在于，由于 $f(i) = i^k$ 是完全积性函数，我们只需提出一个质数，而不需考虑其指数。
所有要求的的 $g(x, j)$ 中的 x 都是 n 除以一个数下取整得到的，因此 x 不超过 $O(\sqrt{n})$ 个。
求 $g(n, j)$ 的复杂度是 $O\left(\frac{\sqrt{n}}{\log n}\right)$ 的，由质数密度可知。
类似杜教筛的不线性筛退化版，积分一下这一部分复杂度 $O\left(\frac{n^{0.75}}{\log n}\right)$。
由下面的内容可知我们只需得到 $g(\ldots, |\mathrm{prime}|)$，滚动数组后空间 $O(\sqrt{n})$。

Min_25 筛

step 3：求答案的 s 数组
定义 $s(n,j)$ 表示最小质因子 $> p_j$ 的数函数值之和，
$s(n,j) = \sum_{i=1}^{n} [\text{minp}(i) > p_j] f(i)$。
分成质数和合数来计算：

$$s(n,j) = \sum_k c_k \cdot g_k(n, |\text{prime}(N)|) - \sum_{i=1}^{j} f(p_i)$$

$$+ \sum_{k>j, p_k \le \sqrt{n}} \sum_{e>0, p_k^e \le n} f(p_k^e) \cdot \left(s\left(\left\lfloor \frac{n}{p_k^e} \right\rfloor, k \right) + [e > 1] \right)$$

N 表示求和的上界。前面两项是两个前缀和相减，表示 $> p_j$ 的质数的函数值之和，第一项的 Σ 表示 $f(p^c)$ 是多项式，第二项是线性筛预处理好的。
第三项是枚举合数的最小质因子和次数 p^e，要求合数剩下部分最小素因子 $> p$，在 $e \neq 1$ 的时候剩下部分可以是 1，$e = 1$ 时由于质数已经统计所以剩下部分不能为 1。

Min_25 筛

最后的答案就是了 $s(n,0)+1$ 了，其中 1 表示 $f(1)$。递归下去算无需记忆化，复杂度未知，常数很小。
小 trick：如何给 n/i 下取整离散化？$i \leq \sqrt{n}$ 和 $n/i \leq \sqrt{n}$ 时分别存在 i 和 n/i 即可。

[模板]Min_25 筛

定义积性函数 $f(x)$，$f(p^k) = p^k(p^k - 1), p \in \text{prime}$。

求 $\left(\sum_{i=1}^{n} f(i)\right) \bmod 10^9 + 7$，$n \leq 10^{10}$。

[模板]Min_25 筛 Solution

把上面的模板套进来就行了。

$\sum_k c_k \cdot g_k(n, |\text{prime}(N)|)$ 中是加上 $k = 2$ 的 g 乘以 1，加上 $k = 1$ 的 g 乘以 -1。

LOJ6053. 简单的函数

已知 f 是个积性函数，且满足 $f(1)=1, f(p^c)=p\oplus c$，其中 p 是质数，$\oplus$ 表示二进制按位异或。

求 $\left(\sum_{i=1}^{n} f(i)\right) \bmod 10^9+7$，$n\leq 10^{10}$。

LOJ6053. 简单的函数 Solution

思路与模板类似，我们注意到
$f(2)=3, f(p)=p-1\ (p\in \mathrm{prime}, \mathrm{p}\neq 2)$，$f(p^k)$ 也很好算。
于是构造函数：

$$g(n,j)=\sum_{i=1}^{n}[i\in \mathrm{prime} \vee \mathrm{minp}(i)>p_j]i$$

$$h(n,j)=\sum_{i=1}^{n}[i\in \mathrm{prime} \vee \mathrm{minp}(i)>p_j]1$$

$$S(n,j)=\sum_{i=1}^{n}[\mathrm{minp}(i)>p_j]f(i)$$

剩下的转移和递归求解和模板题很类似。

hdu7217

定义一个长度为 n 的整数序列 a 是好的，当且仅当 $\forall 1 \le i < n, a_i \mid a_{i+1}$。
每个询问给定 n, m，求长度为 n，$a_n \le m$ 的好的序列的个数，对 10^9+7 取模。
$T \le 100, 1 \le n, m \le 10^9$, $\max(n, m) > 10^6$ 的组数不超过 10。

hdu7217 Solution

考虑对于一个给定的结尾 a_n，我们如何求解方案数 $f(a_n)$。
把 a_n 分解为 $a_n = p_1^{c_1} p_2^{c_2} \cdots p_k^{c_k}$，显然每个质数的方案数是无关的，那么我们把每个质数的方案数乘起来就好了，且 $f(n)$ 是个积性函数。
那么 $f(p)$ 即为 n，$f(p^k)$ 可以用插板求出，$f(p^k) = \binom{n+k-1}{k-1}$。
最后筛一下即可。