数学 08

- 整除分块
- 积性函数与数论函数
- 狄利克雷卷积及其性质
- 题目选讲

整除分块

整除分块用来在 $O(\sqrt{n})$ 时间计算一类整除问题，是后续数论函数计算的基础

问题引入

考虑这样一个问题，给定 n，如何计算

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

最简单的办法是一次遍历，时间复杂度 $O(n)$

```
int ans = 0;
for (int i = 1; i <= n; ++i)
    ans += n / i;
```

打表带来的发现

但是打表发现，对于连续的 i，整除值 $\left\lfloor \frac{n}{i} \right\rfloor$ 趋于相同

例如 $n = 10$，有如下关系

```
|-----------------|----------|
| i               | n / i    |
| --------------- | -------  |
| 1               | 10       |
| 2               | 5        |
| 3               | 3        |
| 4, 5            | 2        |
| 6, 7, 8, 9, 10  | 1        |
| --------------- | -------  |
```

具体来讲，设左端点为 l，则右端点

$$r = \left\lfloor \frac{n}{\left\lfloor \frac{n}{l} \right\rfloor} \right\rfloor$$

证明这个结论

考虑证明，对于给定的 l，计算满足 $\left\lfloor \frac{n}{l} \right\rfloor = \left\lfloor \frac{n}{r} \right\rfloor$ 的正整数 r 的上界

首先有

$$\left\lfloor \frac{n}{l} \right\rfloor = \left\lfloor \frac{n}{r} \right\rfloor \leq \frac{n}{r}$$

根据分式性质

$$\frac{r}{n} \leq \frac{1}{\left\lfloor \frac{n}{l} \right\rfloor}$$

于是

$$r \leq \frac{n}{\left\lfloor \frac{n}{l} \right\rfloor}$$

由于 r 是正整数，于是得到他的上界为

$$\left\lfloor \frac{n}{\left\lfloor \frac{n}{l} \right\rfloor} \right\rfloor$$

代码演绎与复杂度分析

回到一开始的问题，对应的代码为

```
for (int l = 1, r; l <= n; l = r + 1) {
    r = n / (n / l);
    ans += (n / l) * (r - l + 1);
}
```

总的分块段数不超过 $2\sqrt{n}$，因此整除分块总的时间复杂度为 $O(\sqrt{n})$

数论函数

所谓数论函数，说的是这样的一个映射

$$f : \mathbb{Z}^{+} \to \mathbb{C}$$

通俗来讲，定义域为正整数，值域为复数的函数，我们称之为数论函数

积性函数

对于一个数论函数 $f(n)$，如果 $f(1) = 1$，并且选取任意两个互质的正整数 a，b，满足 $f(a \cdot b) = f(a) \cdot f(b)$，那么我们称之为积性函数

如果不考虑 a，b 互质，仍然有 $f(a \cdot b) = f(a) \cdot f(b)$，那么我们称之为完全积性函数

常见的数论函数以及相应的积性如下所示

- 单位函数 $\epsilon(n)=\begin{cases}1, & n=1\\ 0, & n\neq 1\end{cases}$，完全积性
- 幂函数 $\mathrm{Id}_k(n)=n^k$，当 $k=1$ 时，为恒等函数 $\mathrm{Id}(n)$，完全积性
- 除数函数 $\sigma_k(n)=\sum\limits_{d|n} d^k$，当 $k=1$ 时为因数和 $\sigma(n)$ 函数，当 $k=0$ 时为因数个数函数 $d(n)$，积性
- 欧拉函数 $\phi(n)$，积性

线性筛遇上积性函数

回顾一下线性筛法，我们确保每个合数都会被其最小的质因子筛去，这个筛去的过程一定会发生，并且只发生一遍，由此我们确保了线性筛 $O(n)$ 的时间复杂度

线性筛遇上积性函数，可以帮助我们在 $O(n)$ 时间复杂度内计算区间内所有积性函数的值

线性筛欧拉函数

我们知道，在欧拉筛中，合数 n 被其最小的质因子 p 筛去

设 $n = n' \cdot p$

若 $p \mid n'$，则 n' 含有 n 的所有质因子

所以

$$
\begin{aligned}
\phi(n) &= n \cdot \prod_{i=1}^{n}(1 - \frac{1}{p_i}) \\
&= p \cdot n' \cdot \prod_{i=1}^{n}(1 - \frac{1}{p_i}) \\
&= p \cdot \phi(n')
\end{aligned}
$$

若 $p \nmid n'$，则 $(p,\ n') = 1$，由积性函数性质得到

$$\phi(n) = \phi(p) \cdot \phi(n') = (p-1) \cdot \phi(n')$$

```
for(int i = 2; i <= n; ++i){
    if(vis[i]) prime[++cnt] = i, phi[i] = i - 1; // 素数的欧拉函数值为 i - 1
    for(int j = 1; j <= cnt && prime[j] * i <= n; ++j){
        vis[prime[j] * i] = 0;
        phi[prime[j] * i] = phi[i] * (prime[j] - 1);
        if(i % prime[j] == 0) {
            phi[prime[j] * i] = prime[j] * phi[i];
            break;
        }
    }
}
```

线性筛因数个数

设 n 有 k 个质因子，每个质因子出现的次数为 a_i，则 n 的因数个数为

$$d(n) = \prod_{i=1}^{k}(1 + a_i)$$

设 p 为 n 的最小质因子，有 $n = p \cdot n'$，用 num_i 表示 i 的最小质因子个数

若 $p \mid n'$，即 p 也是 n' 的最小质因子

若

$$d(n') = (1 + a_1)(1 + a_2) \cdots (1 + a_k)$$

则

$$d(n) = (2 + a_1)(1 + a_2) \cdots (1 + a_k)$$

因此

$$d(n) = \frac{d(n')}{(1 + num_{n'})} \cdot (2 + num_{n'}),\ num_n = num_{n'} + 1$$

若 $p \nmid n'$，根据 $d(n)$ 的积性函数性质，有

$$d(n) = d(n') \cdot d(p),\ num_n = 1$$

转化为代码如下所示

```
for (int i = 1; i <= n; ++i) {
    if (!vis[i]) prime[++cnt] = i, num[i] = 1, d[i] = 2;
    for (int j = 1; j <= cnt && i * prime[j] <= n; ++j) {
        vis[i * prime[j]] = 1;
        d[i * prime[j]] = d[i] * d[prime[j]];
        num[i * prime[j]] = 1; // i * prime[j] 中第一次出现最小质因子 prime[j]
        if (i % prime[j] == 0) {
            d[i * prime[j]] = d[i] / (num[i] + 1) * (num[i] + 2);
            num[i * prime[j]] = num[i] + 1; // i * prime[j] 中再一次出现最小质因子 prime[j]
            break;
        }
    }
}
```

狄利克雷卷积

狄利克雷卷积是一种定义在数论函数上的二元运算

我们将卷积记作 $*$，与普通的乘积 $\cdot$ 作区分，那么狄利克雷卷积的形式如下

$$(f * g)(n) = \sum_{d|n} f(d)g(\frac{n}{d})$$

也写作

$$(f * g)(n) = \sum_{ab=n} f(a)g(b)$$

狄利克雷卷积性质

运算律

狄利克雷卷积满足如下运算律

- 交换律，即 $f * g = g * f$
- 分配律，即 $(f + g) * h = f * h + g * h$
- 结合律，即 $(f * g) * h = f * (g * h)$

单位元 $\sigma(n)$

我们发现，$\epsilon(n) * f(n) = \sum_{d|n} \epsilon(d) f(\frac{n}{d})$

当且仅当 $d = 1$ 时，$\epsilon(n) = 1$，因此

$$\begin{aligned} \epsilon(n) * f(n) &= \sum_{d|n} \epsilon(d) f(\frac{n}{d}) \\ &= f(n) \end{aligned}$$

所以 $\epsilon(n)$ 是狄利克雷卷积的单位元

狄利克雷卷积的逆

设 $f * g = \epsilon$，则称 g 为 f 的狄利克雷卷积逆元，记为 $g = f^{-1}$

下面取特殊值研究一下 g 的性质

取 $n=1$，则 $(f*g)(1)=\sum\limits_{d|1} f(d)g(\frac{1}{d})=f(1)g(1)=\epsilon(1)=1$

，于是 $g=\frac{1}{f(1)}$，这说明，$f(1)\neq 0$ 是 g 存在的必要条件

取 $n=2$，则 $(f*g)(2)=\sum\limits_{d|2} f(d)g(\frac{1}{d})=f(1)g(2)+$

$f(2)g(1)=\epsilon(0)=0$，于是

$$g(2)=-\frac{f(2)g(1)}{f(1)}$$

取 $n=3$，则 $(f*g)(3)=\sum\limits_{d|3} f(d)g(\frac{1}{d})=f(1)g(3)+f(3)g(1)=\epsilon(0)=0$，于是

$$g(3)=-\frac{f(3)g(1)}{f(1)}$$

取 $n=4$，则 $(f*g)(4)=\sum\limits_{d|4} f(d)g(\frac{1}{d})=f(1)g(4)+f(2)g(2)+f(4)g(1)=\epsilon(0)=0$，于是

$$g(4)=-\frac{f(4)g(1)+f(2)g(2)}{f(1)}$$

我们归纳

当 $n = 1$ 时

$$g(n) = \frac{1}{f(1)}$$

当 $n > 1$ 时

$$g(n) = -\frac{\sum\limits_{d|n,\ d>1} f(d)g(\frac{n}{d})}{f(1)}$$

代入证明

当 $n=1$ 时，$(f*g)(n)=\sum_{d|1} f(d)g(\frac{1}{d})=f(1)g(1)=1=\epsilon(n)$

当 $n > 1$ 时

$$
\begin{aligned}
(f * g)(n) &= \sum_{d|n} f(d)g(\frac{n}{d}) \\
&= f(1)g(n) + \sum_{d|n,\ d>1} f(d)g(\frac{n}{d}) \\
&= -\frac{f(1)}{f(1)} \sum_{d|n,\ d>1} f(d)g(\frac{n}{d}) + \sum_{d|n,\ d>1} f(d)g(\frac{n}{d}) \\
&= 0 \\
&= \epsilon(n)
\end{aligned}
$$

综合上述，$(f * g)(n) = \epsilon(n)$

因此，我们上面归纳出的 g 是 f 的狄利克雷卷积逆元

从抽象代数角度看（选讲）

我们设所有的数论函数集合为 S，狄利克雷卷积为集合上的乘法运算，函数加法为集合上的加法运算

前面得到了狄利克雷卷积的结合律，乘法分配律，交换律，乘法单位元，因此 $< S, *, + >$ 构成 **整环**

但由于并不是每一个数论函数 f 都满足 $f(1) = 1$，所以不是每一个数论函数 f 都存在狄利克雷卷积逆元，因此不构成 **域**

积性函数的狄利克雷卷积还是积性函数（证明选讲）

我们需要证明

$$(f*g)(a\cdot b)=(f*g)(a)\cdot(f*g)(b)$$

考虑展开右边的表达式，有

$$(f*g)(a)\cdot(f*g)(b)=\sum_{d_1|a}f(d_1)g(\frac{a}{d_1})\cdot\sum_{d_2|b}f(d_2)g(\frac{a}{d_2})$$

注意到 $(a,\ b)=1$，即 $a,\ b$ 互质

因此，当 d_1，d_2 分别遍历 a，b 的因子时，$d_1 \cdot d_2$ **遍历且仅遍历** ab 的所有因子**一次**

举个例子来理解，若 $a = 3$，$b = 14$，那么 $d_1 = 1$，3，$d_2 = 1$，2，7，我们发现 $d_1 d_2$ 将遍历 1，2，3，7，6，21 一次

但是，当 $a = 2$，$b = 14$ 时，$d_1 = 1$，2，$d_2 = 1$，2，7，我们发现 $d_1 d_2$ 将遍历 14 两次，分别由 $1 \cdot 14$ 和 $2 \cdot 7$ 构成

严格来讲，假设 $x_1, x_2 \mid a, y_1, y_2 \mid b$，有 $x_1y_1 = x_2y_2$，那么一定有

$$min\{x_i, y_j\} \mid max\{x_i, y_j\}, i, j \in \{1, 2\}$$

所以 a, b 互质的时候 $d_1 \cdot d_2$ 只遍历 ab 的所有因子一次

这样的意义在于，我们可以把展开的表达式继续化简

$$
\begin{aligned}
(f*g)(a)\cdot(f*g)(b) &= \sum_{d_1|a,\ d_2|b} f(d_1)g(\frac{a}{d_1})f(d_2)g(\frac{b}{d_2}) \\
&= \sum_{d_1|a,\ d_2|b} f(d_1\cdot d_2)g(\frac{a}{d_1}\cdot\frac{b}{d_2}) \\
&= \sum_{d|ab} f(d)g(\frac{ab}{d}) \\
&= (f*g)(ab)
\end{aligned}
$$

因此狄利克雷卷积的积性得到证明，以后就可以大胆使用这个结论了

积性函数的狄利克雷卷积逆元还是积性函数

首先，积性函数 f 满足 $f(1) = 1$，因此积性函数一定存在狄利克雷卷积逆元

同时，积性函数的狄利克雷卷积逆元也是积性函数，证明过于复杂，省略不表

2631. 整除分块

给定正整数 n，计算

$$\sum_{i=1}^{n}\left\lfloor \frac{n}{i} \right\rfloor \ mod\ 998244353$$

数据规定

$30\% : 1 \leq n \leq 10^5$

$60\% : 1 \leq n \leq 10^8$

$100\% : 1 \leq n \leq 10^{12}$

题解

整除分块模板题

2630. 因数个数筛

现在给定正整数 n，希望计算

$$\sum_{i=1}^{n} d(i)\ mod\ 998244353$$

和

$$\sum_{i=1}^{n} i \cdot d(i)\ mod\ 998244353$$

数据规定

$$1 \leq n \leq 10,000,000$$

题解

模板题，线性筛因数个数即可

2632 区间因数和

我们用 $\sigma(n)$ 表示 n 的所有因数的和

例如 $\sigma(6) = 1 + 2 + 3 + 6 = 12$

例如 $\sigma(10) = 1 + 2 + 5 + 10 = 18$

现在给定 L，R，计算

$$\sum_{i=L}^{R} \sigma(i)$$

数据规定

$30\% : 1 \le L < R \le 10^5$

$60\% : 1 \le L < R \le 10^7$

$100\% : 1 \le L < R \le 2 \cdot 10^9$

题解

对于数字 i，在 $1 \sim n$ 中共有 $\left\lfloor \frac{n}{i} \right\rfloor$ 个数可以被 i 整除，因此 i 作为因数一共出现 $\left\lfloor \frac{n}{i} \right\rfloor$ 次

因此只需要处理出

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

之后再做差分即可

考虑整除分块，连续的 i 具有相同的

$$\frac{n}{i}$$

并且对于左端点为 l，右端点

$$r = \left\lfloor \frac{n}{\left\lfloor \frac{n}{l} \right\rfloor} \right\rfloor$$

我们可以等差数列计算

$$\frac{(r-l+1)\cdot(l+r)}{2}\cdot\left\lfloor\frac{n}{l}\right\rfloor$$

将每一段的答案累计即可，总的时间复杂度为 $O(\sqrt{n})$