オイラー関数の多重合成を計算するアルゴリズム

Crimson Global Academy 梶田 光

概要

オイラー関数 $\varphi(n)$ の研究では, オイラー関数を含む方程式の解を研究する.
特に最近はオイラー関数の多重合成 $\varphi^k(n) \coloneqq \underbrace{\varphi(...\varphi}_{k \text{ times}}(n)...)$ についての研究が進み, これを含む方程式の解を数値計算によって求めたい場面が多い.
しかし, $\varphi^k(n)$ は乗法的関数ではないので, 一般的に知られている区間篩をそのまま適用することができない.
そこで今回, 3 つのアイデアを含む新しいアルゴリズムを発見し, Rust で実装した.
以下, 正整数 k, N を固定し, $1 \le n \le N$ のすべての n に対して n の素因数分解と $\varphi(n), \varphi^2(n), ..., \varphi^k(n)$ が(メモリマップを用いて)高速に取得できるような LUT をディスク上に構築するアルゴリズムについて考える.

1. 部分分解とオイラー関数

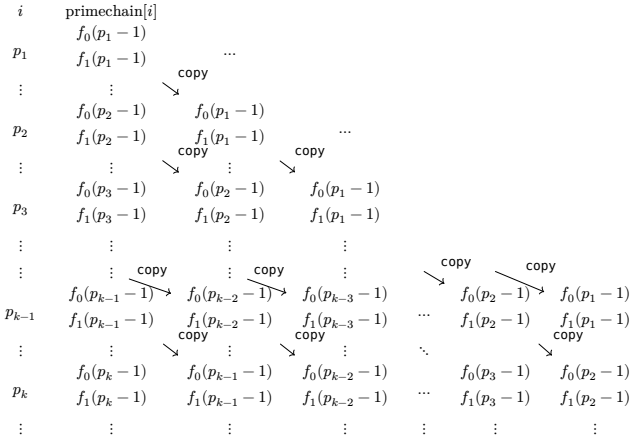
前回の発表では部分分解というアイデアを提案した.
これは正整数 $n \le N$ に対して $n = p_0 p_1 ... p_{m-1}$ $(p_i : \text{prime}, p_i \le p_{i+1}), i_0 = \max\left\{0 \le i \le m \mid \prod_{0 \le j < i} p_j \le \sqrt{N}\right\}, i_1 = \max\left\{i_0 \le i \le m \mid \prod_{i_0 \le j < i} p_j \le \sqrt{N}\right\}$ とおいたとき, $f_0 = \prod_{0 \le j < i_0} p_j, f_1 = \prod_{i_0 \le j < i_1} p_j, f_2 = \frac{n}{f_0 f_1}$ として $n = f_0 f_1 f_2$ という形に書き表す方法である.
これは区間篩内で計算量の大幅な悪化なしに計算することができる.
前回, このような表し方がメモリマップする区間を $O\left(\sqrt{N}\right)$ に抑えつつ素因数分解を取得するのに有用ということを示した.
しかし今回, これが $\varphi^k(n)$ を計算する際にも有用であることがわかった.
一般の n, m について, $d = \gcd(n, m)$ とすると $\varphi(nm) = \varphi(n)\varphi(m)\frac{d}{\varphi(d)}$ という公式がある.
しかしここで, $1 \le \varphi(n) \le n$ が成り立つことはもちろん, $1 \le \varphi(m)\frac{d}{\varphi(d)} \le m$ も成り立ち, さらにこれは自然数である.
したがって, $f([\alpha_0, \alpha_1, \alpha_2]) = \left[\varphi(\alpha_0), \varphi(\alpha_1)\frac{d}{\varphi(d)}, \varphi(\alpha_2)\frac{d'}{\varphi(d')}\right]$ where $d = \gcd(\alpha_0, \alpha_1), d' = \gcd(\alpha_0\alpha_1, \alpha_2)$ のような関数を定義すれば, $\alpha = \alpha_0\alpha_1\alpha_2$ について $f([\alpha_0, \alpha_1, \alpha_2]) = [\alpha'_0, \alpha'_1, \alpha'_2], \alpha' = \alpha'_0\alpha'_1\alpha'_2$ とおいたとき $\varphi(\alpha) = \alpha'$ かつ各 $0 \le i \le 2$ について $\alpha'_i \le \alpha_i$ が成り立つ.
よって $\sqrt{N}$ 以下のすべての n について $\varphi(n)$ を計算しておけば, 任意の $f_0, f_1, f_2 \le \sqrt{N}$ を満たす $n = f_0 f_1 f_2$ について $[f_0, f_1, f_2]$ に先の関数 f を繰り返し適用することによって $\varphi(n), \varphi^2(n), ...$ を計算することができる.

2. $k = 2$ の場合: 調和級数のオーダーと空間計算量の見積り

まず $k = 2$, つまり $\varphi(n), \varphi^2(n)$ までしか計算しなくてよい場合を考える.
$n = f_0(n) f_1(n) f_2(n)$ とおいたとき, $f_0(n), f_1(n), f_2(n) \le \sqrt{N}$ ならば計算できることは先程示した.
$f_0(n), f_1(n) \le \sqrt{N}$ は定義上成り立っているから, $f_2(n) > \sqrt{N}$ の場合を考える.
前回示した定理より, $f_2(n)$ は素数であるから, $\alpha = f_0 f_1 \le \sqrt{N}$ とおくと $\varphi(n) = \varphi(\alpha)(f_2(n) - 1)$.
ここから $\varphi^2(n)$ を計算するには $f_2(n) - 1$ の分解の情報が必要である.
そこで, primechain という長さ N の配列を用意する.
区間篩が $\sqrt{N} < [\text{start}, \text{end}]$ の区間で実行されているとすると, primechain の $\left[\frac{\text{start}}{2}, \frac{\text{end}}{2}\right], \left[\frac{\text{start}}{3}, \frac{\text{end}}{3}\right], \left[\frac{\text{start}}{4}, \frac{\text{end}}{4}\right], ..., \left[\frac{\text{start}}{\sqrt{N}}, \frac{\text{end}}{\sqrt{N}}\right]$ の部分をメモリマップする.
$\text{start} \le n \le \text{end}$ かつ n が素数 (つまり $f_0(n) = f_1(n) = 1$)であれば, $[f_0(n-1), f_1(n-1)]$ を primechain$[n]$ に記録する.
その後, $f_2(n) > \sqrt{N}$ で関数 f の繰り返しによって $\varphi^2(n)$ を計算できない場合, primechain$[f_2(n)]$ から $[f_0(f_2(n)-1), f_1(f_2(n)-1)]$ を取得.
$f_2(f_2(n)-1)$ がそこから計算でき, これが $\sqrt{N}$ 以下なら $\varphi(n)$ は $\varphi(\alpha) \cdot f_0(f_2(n)-1) \cdot f_1(f_2(n)-1) \cdot f_2(f_2(n)-1)$ と $\sqrt{N}$ 以下の正整数の積で表せる.
そうでなければ, $f_2(f_2(n)-1)$ は素数なので $\varphi(n) = \varphi(\varphi(\alpha) \cdot f_0(f_2(n)-1) \cdot f_1(f_2(n)-1)) \cdot (f_2(f_2(n)-1) - 1)$ と計算すればよい.
ここでメモリマップする範囲は調和級数の発散する速度より $O((\text{end} - \text{start}) \log N)$ 程度にしかならず, 区間の大きさを $\sqrt{N}$ 程度にとれば全体を通しての空間計算量は $O\left(\sqrt{N} \log N\right)$ ですむ.

3. 一般の場合: primechain の繋げ方

まず, n の分解 $n = f_0 f_1 f_2$ を考える.
もし $f_2 \le \sqrt{N}$ であれば, $[f_0, f_1, f_2]$ に f を繰り返し適用し続ければよい.
それ以外の場合, $\alpha = f_0 f_1$ とおくと $\alpha = \frac{n}{f_2} < \frac{N}{\sqrt{N}} = \sqrt{N}$ で, $\varphi(n) = \varphi(\alpha f_2) = \varphi(\alpha)(f_2 - 1)$ である.
$\varphi^2(n)$ を計算するためには, $f_2 - 1$ の分解 $f_2 - 1 = f'_0 f'_1 f'_2$ が必要である.
もし $f'_2 \le \sqrt{N}$ であれば, $\varphi(n) = \varphi(\alpha) f'_0 f'_1 f'_2$ は $\sqrt{N}$ 以下の正整数の積で書けるから f を繰り返し適用すればよい.
それ以外の場合, $\alpha' = \varphi(\alpha) f'_0 f'_1$ とおくと $k' = \frac{\varphi(n)}{f'_2} < \frac{n}{\sqrt{N}} \le \frac{N}{\sqrt{N}} = \sqrt{N}$ で, $\varphi^2(n) = \varphi(\alpha')(f'_2 - 1)$.
$\varphi^3(n)$ を計算するためには, $f'_2 - 1$ の分解 $f'_2 - 1 = f''_0 f''_1 f''_2$ が必要である.
もし $f''_2 \le \sqrt{N}$ であれば, $\varphi^2(n) = \varphi(\alpha') f''_0 f''_1 f''_2$ は $\sqrt{N}$ 以下の正整数の積で書けるから f を繰り返し適用すればよい.
それ以外の場合, $\alpha'' = \varphi(\alpha') f''_0 f''_1 f''_2$ とおくと $\alpha'' = \frac{\varphi^2(n)}{f''_2} < \frac{n}{\sqrt{N}} \le \frac{N}{\sqrt{N}} = \sqrt{N}$ で, $\varphi^3(n) = \varphi(\alpha'')(f''_2 - 1)$.
f''' を $f^{(3)}$, f'''' を $f^{(4)}$ などして表記すると, $\varphi^k(n)$ を計算するためには $f_0^{(k-1)}, f_1^{(k-1)}, f_2^{(k-1)}$ までが必要である.
そこで, 順に $f_2, f'_2, ...$ の数列をつなげていき, 以下のように primechain に記録することで $O\left(k\sqrt{N} \log N\right)$ の空間計算量で計算する.
なお, このとき primechain は $N \times (k-1)$ の二次元配列 (要素は 32 ビット符号なし整数のペア)とし, $p_1 > \sqrt{N}, f_2(p_1 - 1) \le \sqrt{N}, f_2(p_2 - 1) = p_1, f_2(p_3 - 1) = p_2, ...$ が成り立っているとする.



全体の時間計算量は $O(kN \log N)$, 空間計算量が $O\left(k\sqrt{N} \log N\right)$, 消費するディスクの容量が $O(kN)$ となる.
一般に $\varphi^k(n) = 1$ となる最小の k は $1 + \log_2 n$ であることが S. S. Pillai [1] によって示されているため, 実用上(オイラー関数の多重合成の性質を調べるという意味で)このアルゴリズムが使われるのは $k \le 1 + \log_2 N$ の範囲のみである.
よってこのアルゴリズムは十分高速である.

最適化について

並列化とメモリマップの領域の最適化についても考察した.
特に後者については, primechain にアクセスするときの添字が素数であることから, wheel sieve の考え方を利用して空間計算量とディスクの容量を $\log \log N$ だけ落とすことができる.

テスト

他に RAM に入らない範囲の n の $\varphi^k(n)$ を計算するアルゴリズムが知られていないので, 既存のアルゴリズムとの速度比較はできない.
しかし, アルゴリズムの正当性のテストのため, RAM に入る範囲の $N = 10^8, k = 4$ で通常のエラトステネスの篩と比較し, $1 \le n \le N$ の範囲で $\varphi(n)$ から $\varphi^4(n)$ までの値がすべて一致することを確認した.

展望

今回のアルゴリズムはかなりオイラー関数固有の性質を活用しているため, 他に多重合成が研究されている乗法的関数 (約数の和関数など) にどこまで応用できるかはまだわかっていない.
また, 同じ問題を解くアルゴリズムで, 時間と空間のトレードオフなしにこのアルゴリズムの計算量を改善することはできないと予想するが, これを解決するのはとても難しいと思われる.

参考文献

[1] S. S. Pillai, “On a function connected with ϕ (n),” Bulletin of the American Mathematical Society, vol. 35, no. 6, pp. 837–841, 1929.