18.785: Analytic Number Theory, MIT, spring 2007 (K.S. Kedlaya)
Brun's combinatorial sieve

In this unit, we describe a more intricate version of the sieve of Eratosthenes, introduced by Viggo Brun in order to study the Goldbach conjecture and the twin prime conjecture. It is most useful for providing lower bounds; for upper bounds, the Selberg sieve (to be introduced in the following unit) is much less painful.

1 Sieve setup

Let $f : \mathbb{N} \to \mathbb{C}$ be an arithmetic function, and suppose we want to estimate the sum of f over primes. More precisely, let P be a set of primes, and put

$$P(z) = \prod_{p \leq z, p \in P} p.$$

If we define

$$\begin{aligned} S(x,z) &= \sum_{n \leq x, (n, P(z))=1} f(n), \\ A_d(x) &= \sum_{n \leq x, n \equiv 0 \pmod{d}} f(n) \end{aligned}$$

(with the dependence on P and f suppressed from the notation), we have

$$S(x,z) = \sum_{d | P(z)} \mu(d) A_d(x).$$

As before, suppose there is a multiplicative function g such that for d squarefree with all prime factors in P,

$$A_d(x) = g(d)X + r_d(x),$$

with $X = X(x)$ independent of d, and the error term $r_d(x)$ small when d is small relative to x (in a sense to be made precise later). Suppose further that

$$g(p) \in [0,1) \quad (p \in P); \qquad g(p) = 0 \quad (p \notin P). \tag{1}$$

(If we need to take $g(p) = 1$, then we cannot expect to get much of a contribution from numbers not divisible by p; we should resign ourselves to this, and instead remove p from P.) Then we can rewrite

$$\begin{aligned} S(x,z) &= V(z)X + R(x,z) \\ V(z) &= \prod_{p | P(z)} (1 - g(p)) \\ R(x,z) &= \sum_{d | P(z)} \mu(d) r_d(x). \end{aligned}$$

If z is small relative to x, which in practice will mean $z < x^\alpha$ for some cutoff $\alpha \in (0,1)$, we may be able to show that the main term $V(z)X$ dominates the error term $R(x,z)$. Again, the main term is what you would predict from the heuristic that if an integer is chosen randomly, its divisibilities by different primes should act like independent random events.

For instance, if P is the set of all primes and $z \geq x^{1/2}$, then $S(x,z) = \sum_{p \leq x} f(p)$. If f is the function

$$f(n) = \begin{cases} 1 & n-2 \text{ prime} \\ 0 & \text{otherwise,} \end{cases}$$

then by the error term in the prime number theorem for arithmetic progressions,

$$r_d(x) = O(x \log^{-A} x)$$

for any fixed $A > 0$. (It is now important that we have that bound uniformly in d!) Also, $S(x, x^{1/2})$ counts twin primes up to x, whereas $S(x, x^{1/(N+1)})$ counts primes p such that $p+2$ has no prime factor less than $x^{1/(N+1)}$, and hence has at most N prime factors.

2 Brun's combinatorial sieve

We would like somewhat finer control than was provided by the sieve of Eratosthenes; the trouble is that $R(x,z)$ has too many terms for us to be able to control it.

Brun's approach to get aronud this is to truncate the Möbius function by restricting it to suitable subsets D^+ and D^-, subject to the restriction that for n a product of primes in P, the incomplete convolutions

$$\delta^+(n) = \sum_{d|n, d \in D^+} \mu(d), \qquad \delta^-(n) = \sum_{d|n, d \in D^-} \mu(d)$$

satisfy

$$\delta^-(n) \leq \delta(n) \leq \delta^+(n) \tag{2}$$

for

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

One such choice would be to take D^+ and D^- to consist of all squarefree numbers whose number of distinct prime factors is even or odd, respectively. This choice is much too crude; we should instead make a choice that allows some cancellation in δ^- and δ^+ without messing up the inequality (2). Moreover, we want to restrict D^+ and D^- to be subsets of $\{1, \ldots, y\}$ for some y which is not too large compared to x.

Let $\lambda^+(d)$ and $\lambda^-(d)$ denote the functions which agree with μ on D^+ and D^-, respectively, and are zero elsewhere. Put

$$\begin{aligned} V^{\pm}(z) &= \sum_{d|P(z)} \lambda^{\pm}(d)g(d) \\ R^{\pm}(x,z) &= \sum_{d|P(z)} \lambda^{\pm}(d)r_d(x). \end{aligned}$$

Then by virtue of (2), we have

$$V^-(z)X + R^-(x,z) \le S(x,z) \le V^+(z)X + R^+(x,z). \tag{3}$$

It is not at all obvious how one can usefully arrange for D^+, D^- to satisfy (2); here is Brun's choice. For d a squarefree positive integer, write $d = p_1 \cdots p_r$ with $p_1 > \cdots > p_r$. Set

$$\begin{aligned} D^+ &= \{d = p_1 \cdots p_r : p_m < y_m \quad m \text{ odd}\} \\ D^- &= \{d = p_1 \cdots p_r : p_m < y_m \quad m \text{ even}\}, \end{aligned}$$

where $y_1, y_2, \dots$ are certain parameters which may depend on d. (By convention, $1 \in D^{\pm}$.) We then have the following.

Lemma 1. *With notation as above, let $V_n(z)$ be the sum of $g(p_1 \cdots p_n)V(p_n)$ over sequences $p_1 > \cdots > p_n$ of primes such that:*

(a) $p_1 < z$;

(b) $p_n \ge y_n$;

(c) $p_m < y_m$ *for* $m < n$ *with* $m \equiv n \pmod 2$.

Then

$$\begin{aligned} V(z) &= V^+(z) - \sum_{n \equiv 1\,(2)} V_n(z) \\ V(z) &= V^-(z) + \sum_{n \equiv 0\,(2)} V_n(z) \end{aligned}$$

and so

$$V^-(z) \le V(z) \le V^+(z). \tag{4}$$

Proof. Exercise. □

In particular, for a given n, we deduce (2) from (4) by rigging up the set P so that $P(z) = n$ and putting $g(d) = 1$ for all d.

The functions λ^+ and λ^- given above are together called the *combinatorial sieve* with parameters $y_1, y_2, \dots$. To use it, one must bound

$$R(x,y) = \sum_{d<y, d|P(z)} |r_d(x)|,$$

for y such that $D^\pm \subset \{1, \dots, y\}$; in this case $R(x,y) \geq |R^\pm(x,z)|$, giving error bounds in (3). One must also bound $V^\pm(z)$.

3 Setting some parameters

To turn this into an actual numerical theorem, we must set the sieve parameters; we do this following Iwaniec-Kowalski. Remember that we may allow the y_i to depend on d.

Write $d = p_1 \cdots p_r$ with $p_1 > \cdots > p_r$; we now take

$$y_m = (y/(p_1 \cdots p_m))^{1/\beta},$$

where $\beta > 1$ will be specified later. This makes it clear that all elements of $D^+ \cup D^-$ belong to $\{1, \dots, y\}$ except possibly for single primes in D^-. We can remedy this by requiring $z \leq y$; more precisely, we will take $z = y^{1/s}$ for some $s \geq \beta$.

We will also need to make some restriction on the multiplicative function g. Namely, we assume that for some $K > 1$ and $\kappa > 0$, we have for all w, z,

$$\prod_{w \leq p < z} (1 - g(p))^{-1} \leq K \left(\frac{\log z}{\log w} \right)^\kappa. \tag{5}$$

We refer to κ as a *sieve dimension* of the function g. This number is quite critical; it will determine how large we can make z compared to y, which determines how many small primes we can use for sieving.

4 Bounding the main term

We need an upper bound on $V^+(z)$ and a lower bound on $V^-(z)$; we get both of these by getting an upper bound on $V_n(z)$. First, let us simplify the sum by relaxing the summation conditions. We claim that for any tuple $p_1, \dots, p_n$ appearing in the sum defining $V_n(z)$, and any $m < n$,

$$p_1 \cdots p_{m-1} p_m^\beta < y. \tag{6}$$

Namely, if $m \equiv n \pmod 2$, we have the stronger inequality

$$p_1 \cdots p_{m-1} p_m^{1+\beta} < y.$$

If $m > 1$ and $m \not\equiv n \pmod 2$, we have

$$p_1 \cdots p_{m-1} p_m^{\beta} < p_1 \cdots p_{m-2} p_{m-1}^{1+\beta} < y.$$

Finally, if $m = 1$ and $m \not\equiv n \pmod 2$, we have

$$p_1^{\beta} < z^{\beta} = y^{\beta/s} \leq y.$$

From (6), we deduce by induction on m that

$$p_1 \cdots p_m < y^{1-(1-\beta^{-1})^m} \qquad (m = 1, \dots, n-1).$$

In particular,

$$p_n \geq (y/(p_1 \cdots p_{n-1}))^{1/(\beta+1)} \geq y^{\frac{1}{\beta+1}(1-\beta^{-1})^{n-1}} \geq y^{\frac{1}{\beta}(1-\beta^{-1})^{n}} \geq z_n$$

if we put

$$z_n = z^{(1-\beta^{-1})^n}.$$

We will now retain only the conditions $z > p_1 > \cdots > p_n \geq z_n$ on the primes, which will make the sum bigger because every summand is nonnegative. That is,

$$\begin{aligned} V_n(z) &\leq \sum_{z > p_1 > \cdots > p_n \geq z_n} g(p_1 \cdots p_n) V(p_n) \\ &\leq \frac{1}{n!} V(z_n) \left(\sum_{z_n \leq p < z} g(p) \right)^n \\ &\leq \frac{1}{n!} V(z_n) \left(\log \frac{V(z_n)}{V(z)} \right)^n . \end{aligned}$$

Here is where we need the assumption (5) about the sieve dimension. It implies

$$\frac{V(z_n)}{V(z)} \leq K(1 + (\beta - 1)^{-1})^{\kappa n} < K e^{n/b}$$

for $\beta = \kappa b + 1$ (using the bound $1 + x \leq e^x$ for $x = (\beta - 1)^{-1} = 1/(\kappa b)$), which gives us

$$\begin{aligned} V_n(z) &< \frac{K}{n!} \left(\frac{n}{b} + \log K \right)^n e^{n/b} V(z) \\ &\leq \frac{K}{n!} \left(\frac{n}{b} e^{1/b} \right)^n K^b V(z) \end{aligned}$$

(using the bound $1 + x \leq e^x$ for $x = b(\log K)/n$). Since $n! \geq e(n/e)^n$ (by taking logs and comparing integrals), we obtain

$$V_n(z) < e^{-1} a^n K^{b+1} V(z)$$

for $a = b^{-1}e^{1+b^{-1}}$.

To conclude, we clean things up a bit. Remember that we were at liberty to choose $\beta > 1$, which is equivalent to choosing $b > 0$. By taking b sufficiently large, we can force $a < 1$; for instance, we could take $b = 9$ to get $a < e^{-1}$. Note also that because

$$p_1 > \cdots > p_n \geq y_n = (y/(p_1 \cdots p_n))^{1/\beta},$$

we have $p_1^{n+\beta} > y$. Since we also have $p_1 < z = y^{1/s}$, we deduce that $V_n(z) = 0$ unless $n + \beta > s$. Therefore

$$\sum_{n>0} V_n(z) = \sum_{n>s-\beta} V_n(z) < \frac{a^{s-\beta}}{e(1-a)} K^{b+1} V(z).$$

To conclude, we have the following bound (Theorem 6.1 in Iwaniec-Kowalski).

Theorem 2. *In the combinatorial sieve with parameters $y_1, y_2, \ldots$ as above, and $\beta = 9\kappa+1$, for any multiplicative function $g(d)$ satisfying* (1) *and* (5) *for a given K, and any $s \geq \beta$, for $z = y^{1/s}$ we have*

$$V^+(z) < (1 + e^{\beta-s}K^{10})V(z)$$
$$V^-(z) > (1 + e^{\beta-s}K^{10})V(z).$$

Consequently,

$$(1 - e^{\beta-s}K^{10})V(z)X - R(x, z^s) \leq S(x,z) \leq (1 + e^{\beta-s}K^{10})V(z)X + R(x, z^s).$$

5 Consequences for twin almost-primes

Again consider the example

$$f(n) = \begin{cases} 1 & n-2 \text{ prime} \\ 0 & \text{otherwise.} \end{cases}$$

By applying the combinatorial sieve, we may deduce the following (see exercises).

Theorem 3. *There are infinitely many primes p such that $p + 2$ is the product of at most twenty distinct primes.*

By refinements of the sieving method, Chen was able to prove the following.

Theorem 4. *There are infinitely many primes p such that $p + 2$ is the product of at most two distinct primes.*

This is tantalizingly close to the twin prime conjecture, but it seems that sieving methods fall short of delivering that particular prize.

One can also use the combinatorial sieve to deduce that the number of twin primes $\leq x$ is $O(x/\log^2 x)$; however, since this is a question about an upper bound rather than a lower bound, we will be able to derive this much less painfully using the Selberg sieve.

Exercises

1. Prove Lemma 1. (Hint: use the identity

$$V(z) = 1 - \sum_{p<z} g(p)V(p)$$

plus inclusion-exclusion.)

2. Apply the combinatorial sieve to show that the number of integers less than or equal to x with no prime factors less than $x^{1/20}$ is at least $cx/\log^2 x$ for some $c > 0$. (You will need the prime number theorem in arithmetic progressions with error term, in order to control the error term $R(x, z)$.) Then deduce Theorem 3.