THE SELBERG SIEVE

ARJUN VENKATRAMAN

1. Motivation for Sieves and Sieve Theory

The purpose of sieving and sieve theory is to find approximations and bounds on prime numbers or structures bearing a resemblance to prime numbers. Essentially, given a small set of primes, we want an approximation of how many numbers in a much larger set divide none of those primes. Since all primes in the larger set must satisfy this criterion, this gives us a bound on the number of primes up to x and how fast it grows. Similarly, these methods can be expanded to non-integer sets — if each prime corresponds to a subset, then we try to search for all elements not contained in any such subset.

The most basic sieve is the Sieve of Eratosthenes, named after the Greek mathematician who invented it. The sieve is constructed by taking the first n positive integers and crossing out the number 1. Then the next unmarked number, 2, is a prime, and we cross out all multiples of 2. The next unmarked is 3, which we mark as prime and cross out its multiples. Next is 5, then 7, then 11, and so on. While the sieve is basic, certain ways of rewriting the method allow for very useful conclusions. A modification of one such method is what creates the Selberg sieve.

2. The Selberg Sieve

Defining $\Phi(x, z)$ as the number of integers below x dividing no primes below z, we can use inclusion-exclusion to derive the identity

$$\Phi(x, z) = \sum_{d|P_z} \mu(d) \sum_{d|n\leq x} 1 = \sum_{n\leq x} \sum_{d|n,P_z} \mu(d)$$

where P_z is defined as the product of all primes up to z.

This identity is Legendre's representation of the Sieve of Eratosthenes, coming from the fact that the Möbius function can be summed to yield an inclusion-exclusion counting by prime factors. Setting $z = k$ and maipulating the result by relating $\pi(x)$ and $\Phi(x)$ gives the bound $\pi(x) = O(\frac{x}{\log\log x})$. Using further manipulations, we derive a slightly better bound: $\pi(x) = O(\frac{x\log\log x}{\log x})$.

The goal of Selberg's method, then, is to further reduce this bound by replacing the Möbius function in the Legendre identity with a series of weights set specifically to minimize the bound. Specifically, since the sum of the Möbius function applied to all divisors of k is 1 if $k = 1$ and 0 otherwise, we can set a series of weights λ_n — as long as $\lambda_1 = 1$, we can say that

$$\sum_{d|k} \mu(d) \leq \left(\sum_{d|k} \lambda_d\right)^2.$$

Date: June 10, 2019.

This is the key observation to the creation of this sieve. We can substitute this into Legendre's identity to get

$$\Phi(x,z) \leq \sum_{n\leq x}\left(\sum_{d|(n,P_z)}\lambda_d\right)^2.$$

Splitting up the square sum and rearranging the summation order yields

$$\Phi(x,z) \leq \sum_{d_1,d_2|P_z}\lambda_{d_1}\lambda_{d_2}\sum_{[d_1,d_2]|n\leq x}1$$

where $[m,n]$ is the least common multiple of m and n.

The second summation in this form is an alternative formulation of the function $\lfloor\frac{x}{[d_1,d_2]}\rfloor$. Since a floor is never more than one away from the initial value, this can also be rewritten as $\frac{x}{[d_1,d_2]}+O(1)$. Replacing this in the above expression and moving it to the "outside" of the sum gives

$$\Phi(x,z) \leq x\sum_{d_1,d_2|P_z}\frac{\lambda_{d_2}\lambda_{d_1}}{[d_1,d_2]}+O\left(\sum_{d_1,d_2|P_z}|\lambda_{d_1}||\lambda_{d_2}|\right).$$

What we now desire is to set values of λ_n to minimize both the concrete term and the error term here.

The first choice we make is to replace the summations of divisors of P_z into something more convenient. To so this, we set $\lambda_d = 0$ for $d > z$, leaving us with the nicer

$$\Phi(x,z) \leq x\sum_{d_1,d_2\leq z}\frac{\lambda_{d_2}\lambda_{d_1}}{[d_1,d_2]}+O\left(\sum_{d_1,d_2\leq z}|\lambda_{d_1}||\lambda_{d_2}|\right).$$

Now, isolating the main term, we can multiply out (d_1,d_2) to get

$$\sum_{d_1,d_2\leq z}\frac{\lambda_{d_2}\lambda_{d_1}}{d_1d_2}(d_1,d_2).$$

Next, we use the fact that the sum of the Euler function over the divisors of d is equal to d to get

$$\sum_{d_1,d_2\leq z}\frac{\lambda_{d_2}\lambda_{d_1}}{d_1d_2}\sum_{k|(d_1,d_2)}\phi(k).$$

We can swap the summation order of this expression and combine the dual summation into a square to obtain

$$\sum_{k\leq z}\phi(k)\left(\sum_{k|d\leq z}\frac{\lambda_d}{d}\right)^2.$$

This form can be more easily minimized. Evaluating from the conditions set for λ_n eventually yields a minimum value of $\frac{1}{V(z)}$, where

$$V(z) = \sum_{d\leq z}\frac{\mu^2(d)}{\phi(d)}.$$

This occurs when we set

$$\lambda_k = k \sum_{kmidd \leq z} \frac{\mu(d/k)\mu(d)}{\phi(d)V(z)}.$$

Thus, we have minimized our non-O term to $\frac{x}{V(z)}$.

Next, we turn to the error term. By examining closely the values of λ_n (specifically expanding out $V(z)\lambda_n$), we find that $|\lambda_n| \leq 1$ for all n. Thus

$$\sum_{d_1,d_2 \leq z} |\lambda_{d_1}||\lambda_{d_2}| \leq z^2,$$

giving us a final bound of

Theorem 2.1.

$$\Phi(x,z) \leq \frac{x}{V(z)} + O(z^2).$$

This can be expanded to the general sieve case. In the general case, we have have a set A of elements and a set P of primes and for each prime $p \in P$ we have a set $A_p \subset A$. For all other positive integers d we define A_d as the intersection of the A_p's for all primes that divide it, and we set $A_1 = 1$. We seek to find bounds on $S(A, P, z)$,the number of elements of A which are not contained in any A_p for some $p \mid P(z)$.

For the general Selberg sieve, we take some multiplicative function f and some positive real number X such that the number of elements of A_d can be written as $\frac{X}{f(d)} + R_d$ where R_d is some real number. The Selberg sieve bound states that:

Theorem 2.2. *(Selberg sieve)*

$$S(A, \mathcal{P}, z) \leq \frac{|A|}{V(z)} + O\left(\sum_{d_1,d_2 \leq z; d_1,d_2 | P(z)} |R_{[d_1,d_2]}| \right),$$

where

$$V(z) = \sum_{d \leq z; d|P(z)} \frac{\mu^2(d)}{f_1(d)}$$

and f_1 is the Möbius inversion $\sum_{d|n} \mu(d) f(\frac{n}{d})$.

The proof of this bound is slightly more complicated, but follows the same general steps of the above proof of the sieve bound on $\Phi(x, z)$.

The error term here does not simplify quite as well due to the condition of all the terms of its summation being smaller than 1 is not necessarily met.

3. Counting Primes with the Selberg Sieve

We can now use the Selberg sieve bound for $\Phi(x, z)$ to get a better bound on $\pi(x)$. First, we assume $z \leq x$ to get $\pi(x) \leq \Phi(x, z) + z$. Examining the value of $V(z)$, we get the lower bound $V(z) >> \log z$. From this (and since $z = O(z^2)$) we achieve the upper bound

$$\pi(x) < \frac{x}{\log z} + O(z^2).$$

If we set $z = \sqrt{\frac{x}{\log x}}$, we get:

$$\pi(x) < \frac{x}{1/2(\log x - \log\log x)} + O\left(\frac{x}{\log x}\right)$$
$$\pi(x) < O\left(\frac{x}{\log x}\right) + O\left(\frac{x}{\log x}\right),$$

which yields

Theorem 3.1.
$$\pi(x) = O\left(\frac{x}{\log x}\right)$$

This gives us a stronger bound than can be derived from either the Brun sieve or the Sieve of Eratosthenes.

4. Other Applications

One of the major applications of the Selberg sieve is the proof of the Brun-Titchmarsh theorem. The theorem centers around the question of trying to find the number of primes in an arithmetic progression. Specifically, we want to find a bound on $\pi(x; q, a)$, defined as the number of primes $p \leq x$ congruent to $a \mod k$. The Brun-Titchmarsh theorem, proved using the Selberg sieve, gives a bound as follows:

Theorem 4.1. *(Brun-Titchmarsh) Given positive integers a and k coprime, and x defined such that there exists some $\theta < 1$ for which $k \leq x^\theta$, for any $\epsilon > 0$, there exists some $x_0 > 0$ such that*
$$\pi(x; k, a) \leq \frac{(2+\epsilon)x}{\phi(k)\log(2x/k)}$$
for all $x > x_0$.

This gives us a bound of $\frac{2x}{\phi(k)\log(2x/k)}$ as $x \to \infty$.

Brun-Titchmarsh can in turn be used to prove a number of other results, such as setting the bound of the number of $n \leq x$ such that n and $\phi(n)$ are relatively prime is approximately $\frac{e^{-y}x}{\log\log\log x}$ as $x \to \infty$ (due to Erdös).

In summary, the Selberg sieve — a modification of Legendre's rewritten Sieve of Eratosthenes using modifiable weights instead of the Möbius function that allow for further optimization — can be used to prove bounds on the size of numerous sets of primes or prime-like structures. While some of these bounds can be proven without sieve theory, others are based uniquely on the utilization of this method.

Euler Circle, Palo Alto, CA 94306
E-mail address: arjun.b.venkatraman@gmail.com