SELBERG'S SIEVE - AN INTRODUCTION COURSE NOTES, 2015

ZEÉV RUDNICK

1. The sieve of Eratosthenes

Sieve methods are techniques for estimating sets of primes (or integers) based on restrictions on their divisibility properties, starting from the sieve of Eratosthenes. My goal in this lecture is to explain how to use sieve methods to obtain upper bounds on various prime counting functions.

For instance, let's try to bound the number $\pi(x)$ of primes $p \leq x$, which the PNT says is $\sim x/\log x$. We have previously seen Chebyshev's method, but here we argue differently so will not rely on those results. As substitutes, we will need some very weak bounds on quantities related to the number of primes, much weaker than Chebyshev's theorem.

Lemma 1.1. *i) Let $P(z) := \prod_{p\leq z} p$ be the product of all primes up to z. Then $P(z) \leq z^z$.*

ii)

$$\prod_{p\leq z}(1-\frac{1}{p})^{-1} \gg \log z\ .$$

iii)

$$\sum_{p\leq z}\frac{1}{p} \gg \log\log z\ .$$

Of course, we have seen stronger statements before (Merten's theorems) but the proof we gave for those relied on Chebychev's bounds.

Proof. 1) We trivially have

$$P(z) = \prod_{p\leq z} p \leq \prod_{n\leq z} z = z^z\ .$$

2) Using the geometric series, we have

$$\prod_{p\leq z}(1-\frac{1}{p})^{-1} = \prod_{p\leq z}(1+\frac{1}{p}+\frac{1}{p^2}+\dots) = \sum_{p|n\Rightarrow p\leq z}\frac{1}{n}$$

Date: April 29, 2015.

by unique factorization into primes. Omitting all terms bigger than z leaves us with

$$\prod_{p\leq z}(1-\frac{1}{p})^{-1}\geq \sum_{n\leq z}\frac{1}{n}\sim \log z$$

as claimed.

3) We have, on using $\log(1-y)^{-1}=y+O(y^2)$ for $0<y<1$,

$$\log\prod_{p\leq z}(1-\frac{1}{p})^{-1}=\sum_{p\leq z}\log(1-\frac{1}{p})^{-1}=\sum_{p\leq z}\frac{1}{p}+O(\frac{1}{p^2})=\sum_{p\leq z}\frac{1}{p}+O(1)$$

On the other hand, we saw $\prod_{p\leq z}(1-\frac{1}{p})^{-1}\gg \log z$ which gives

$$\log\prod_{p\leq z}(1-\frac{1}{p})^{-1}\gg \log\log z$$

as required. □

We first estimate $\pi(x)-\pi(z)$ =number of primes $z<p\leq x$, by the number of integers $n\leq x$ which are not divisible by any "small" prime $p\leq z$, that is are co-prime with $P(z):=\prod_{p\leq z}p$:

$$\pi(x)-\pi(z)\leq \mathcal{S}(x,z):=\#\{n\leq x:\gcd(n,P(z))=1\}$$

In fact if $z=\sqrt{x}$ then we would have equality

$$\pi(x)-\pi(\sqrt{x})=\mathcal{S}(x,\sqrt{x})$$

so below we would like to take z as large as $\sqrt{x}$ (but will fail badly!).

We now proceed: We have

$$\mathcal{S}(x,z)=\sum_{n\leq x}\delta(\gcd(n,P(z))$$

where

$$\delta(m)=\begin{cases}1, & m=1\\ 0, & m>1\end{cases}.$$

We use Möbius inversion:

$$\delta(\gcd(a,b))=\sum_{d|\gcd(a,b)}\mu(d)=\sum_{\substack{d|a\\ d|b}}\mu(d)\ .$$

Hence

$$\begin{aligned}\mathcal{S}(x,z)&=\sum_{n\leq x}\sum_{\substack{d|n\\ d|P(z)}}\mu(d)=\sum_{\substack{d|P(z)\\ d\leq x}}\mu(d)\sum_{\substack{n\leq x\\ d|n}}1\\ &=\sum_{\substack{d|P(z)\\ d\leq x}}\frac{x}{d}+O(1)\ .\end{aligned}$$

To proceed further we want to eliminate the condition $d\leq x$ in the sum over d, leaving only the requirement that $d\mid P(z)$. We can achieve this if we

require that z is sufficiently large so that $P(z) \le x$, which requires knowing there are not too many primes (hence might be begging the question, since we are trying to give an upper bound for $\pi(x)$!). We saw that $P(z) \le z^z$, so we choose

$$z \le \frac{\log x}{\log\log x}$$

This implies that

$$\log P(z) \le \log z^z = z\log z \le \frac{\log x}{\log\log x}\log\frac{\log x}{\log\log x} \le \frac{\log x}{\log\log x}\log\frac{\log x}{1} = \log x$$

and hence we are guaranteed that if $z \le \frac{\log x}{\log\log x}$, then $P(z) \le x$. Thus with this restriction on z, we

$$\mathcal{S}(x,z) \le x\sum_{d|P(z)}\frac{\mu(d)}{d} + O(\sum_{d|P(z)} 1)$$

Since we saw that for any multiplicative function α, we have

$$\sum_{d|D}\alpha(d) = \prod_{p|D}\Big(1+\alpha(p)+\cdots+\alpha(p^j)\Big)\,,$$

we have

$$\sum_{d|P(z)}\frac{\mu(d)}{d} = \prod_{p\le z}(1-\frac{1}{p}).$$

Moreover, the number of divisors $d \mid P(z)$ is $2^{\pi(z)} \le 2^z$. Hence

$$\mathcal{S}(x,z) \le x\prod_{\substack{p\le z\\ \text{prime}}}(1-\frac{1}{p}) + O(2^z)$$

Since we have $\pi(x) \le \mathcal{S}(x,z) + \pi(z)$, we obtain

$$\pi(x) \ll \frac{x}{\log z} + 2^z + \pi(z) \ll \frac{x}{\log z} + 2^z\,. \tag{1}$$

We now pick z so as to minimize the RHS of (1) (subject to $z \le \frac{\log x}{\log\log x}$). Because of the remainder term of 2^z we cannot do better than $z = \log x$, in fact should take $z = \log x/\log\log x$, and then we get

$$\pi(x) \ll \frac{x}{\log\log x}\,.$$

This falls far short of the expected answer of $x/\log x$!

Note that even if we could take z of size $\sqrt{x}$, the main term given by Merten's theorem $\prod_{p\le z}(1-\frac{1}{p}) \sim \frac{e^{-\gamma}}{\log z}$ will have the wrong constant $e^{-\gamma} \ne 1$!

2. Selberg's sieve

Recall that

$$\mathcal{S}(x,z)=\sum_{n\leq x}\delta\Big(\gcd(n,P(z))\Big)$$

Selberg's idea was to replace the $\delta\Big(\gcd(n,P(z))\Big)$ by a system of inequalities: Given parameters $\{\lambda_d\}$ such that

- λ_d real
- $\lambda_1=1$

then

$$\delta(m)=\begin{cases}1, & m=1\\ 0, & m>1\end{cases}\Big\}\leq\Big(\sum_{d|m}\lambda_d\Big)^2\,. \tag{2}$$

Indeed, when $m=1$ then both sides are equal, otherwise the LHS is zero while the RHS, being real, is non-negative.

Using (2) to bound the number $\mathcal{S}(x,z)$ of integers $n\leq x$ not divisible by small primes $p\leq z$ gives

$$\mathcal{S}(x,z)=\sum_{n\leq x}\delta\Big(\gcd(n,P(z))\Big)\leq\sum_{n\leq x}\Big(\sum_{\substack{d|n\\ d|P(z)}}\lambda_d\Big)^2\,. \tag{3}$$

The RHS of (3) is a quadratic form in the variables $\{\lambda_d\}$. Selberg succeeded in minimizing it.

To do so, we take λ_d supported on integers $d\leq z$, that is assume $\lambda_d=0$ if $d>z$. Then

$$\begin{aligned}\mathcal{S}(x,z)&\leq\sum_{n\leq x}\Big(\sum_{\substack{d|n\\ d|P(z)}}\lambda_d\Big)^2\\ &=\sum_{\substack{d_1,d_2\leq z\\ \text{squarefree}}}\lambda_{d_1}\lambda_{d_2}\#\{n\leq x: d_1\mid n,\ d_2\mid n\}\end{aligned}$$

Now the two conditions $d_1\mid n$, $d_2\mid n$ are equivalent to n being divisible by the least common multiple $[d_1,d_2]=\mathrm{lcm}(d_1,d_2)$, hence

$$\#\{n\leq x: d_1\mid n, d_2\mid n\}=\#\{n\leq x:[d_1,d_2]\mid n\}=\lfloor\frac{x}{[d_1,d_2]}\rfloor$$

Using $\lfloor y\rfloor=y+O(1)$ gives

$$\mathcal{S}(x,z)\leq x\sum_{d_1,d_2\leq z}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]}+O\Big(\sum_{d_1,d_2}|\lambda_{d_1}||\lambda_{d_2}|\Big) \tag{4}$$

The goal is now to minimize the quadratic form

$$Q(\lambda)=\sum_{\substack{d_1,d_2\leq z\\ \text{squarefree}}}\frac{\lambda_{d_1}\lambda_{d_2}}{[d_1,d_2]}$$

subject to the constraint $\lambda_1 = 1$. Selberg did this in an ingenious elementary fashion, by explicitly diagonalizing Q, obtaining

$$\min_{\substack{d_1,d_2 \leq z \text{ squarefree} \\ \lambda_1 = 1}} Q(\lambda) = \frac{1}{S(z)}, \quad S(z) = \sum_{\substack{d \leq z \\ d \text{ squarefree}}} \frac{1}{\phi(d)}$$

where $\phi(d) = \#\{1 \leq a \leq d : \gcd(a,d) = 1\}$ is Euler's totient function. Along the way he obtained that the minimizing vector has $|\lambda_d| \leq 1$. Inserting into (4) gives

$$\mathcal{S}(x,z) \leq \frac{x}{S(z)} + z^2$$

Note that the remainder term here is z^2 rather than 2^z in the sieve of Eratosthenes. One then shows that

$$S(z) = \sum_{\substack{d \leq z \\ d \text{ squarefree}}} \frac{1}{\phi(d)} \gtrsim \log z \tag{5}$$

to deduce that

$$\mathcal{S}(x,z) \lesssim \frac{x}{\log z} + z^2$$

This is an exponential improvement on the error term in comparison with the upper bound of $x/\log z + 2^z$ in the sieve of Eratostenes (1). We can now write

$$\pi(x) \leq \mathcal{S}(x,z) + \pi(z) \lesssim \frac{x}{\log z} + z^2$$

and pick $z \approx x^{1/2-o(1)}$ to find

$$\pi(x) \lesssim 2\frac{x}{\log x}$$

which is the expected upper bound (up to a constant)!

It remains to show (5), that is that

$$S(z) := \sum_{\substack{d \leq z \\ d \text{ squarefree}}} \frac{1}{\phi(d)} \gtrsim \log z$$

Since $\phi(d) \leq d$, $S(z) \geq \sum_{d \leq z, \text{ squarefree}} 1/d$. Since squarefree integers have a positive density $1/\zeta(2)$, it is easy to see using summation by parts that the latter sum is $\gtrsim \frac{1}{\zeta(2)} \log z$. A more careful analysis of the original sum $S(z)$ gives a constant of 1.