

Generation of Prime Numbers

Mark Anthony Caruana

Introduction

Introduction Prime numbers are by definition numbers which are only divisible by one and by themselves. It can be proved that such numbers are infinite, as are after all, the Real Numbers, or the Natural numbers. In the following pages will be trying to shed some light over the following unanswered question:

”Is it possible to come up with some form of equation with which one can generate such numbers?”

It is interesting to know that ever since Antiquity, mathematicians have always been haunted by this infamous question which, in virtue seems to have no straight forward answer.

History

Mathematicians like Eratosthenes (275-194 B.C., Greece) managed however to come up with interesting methods of how to find prime numbers. One of these methods is known as 'The Sieve of Eratosthenes'. The method is rather crude and time consuming but rather efficient. All one must do is to build up a chart, or a list containing the natural numbers. Then one must simply follow the set of simple instructions listed below:

- Cross out 1, (as one is not prime).

- Circle 2, because it is the smallest positive even prime. Now cross out every multiple of 2 present in the chart.
- Circle 3, the next number in the list, and again, cross out all the multiples of 3.
- Omit 4, (as four has four has been already crossed out).
- Circle 5, etc.

In later years the mathematician Euler came up with an equation which could generate 40 primes, which for the time was a break through. The equation is $y = x^2 + x + 41$ and could generate prime numbers for x between 0 and 39. In recent times, someone, whose identity is not known, came up an equation which could generate 45 prime numbers; this equation is presently the record breaker. The equation is $y = 36x^2 - 810x + 2753$, and can generate primes for x between 0 and 44.

Prime Number generation

There are various algorithms by which prime numbers can be created (one of them is the Sieve of Eratosthenes); however we are after an algebraic equation which can generate such numbers.

Approximations

These equations have not been conceived to created EXACT sequences of prime numbers but rather a sequence of numbers which are approximately prime i.e. such equations will generate numbers such as: 1.99, 2.87, 4.98 . . . These sequences when rounded off very often will give us a very good sequence of prime numbers.

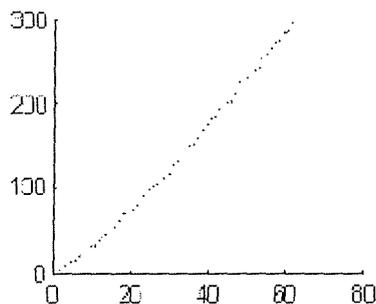
- The Linear Approximation

In this method we will derive a linear equation which will relate the Natural Numbers to the Prime Numbers.

Build a table like the following:

x values	1	2	3	4	5	...
y values	2	3	5	7	11	...

From the above table, the following scatter plot is constructed.



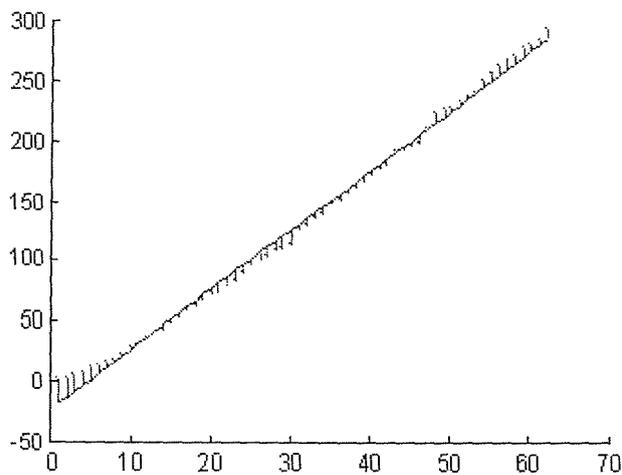
From the above figure and table we can obtain the line of best fit $Y = mx + c$, where

$$m = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

and

$$c = \frac{1}{n} \sum_{i=1}^n y_i - \frac{m}{n} \sum_{i=1}^n x_i$$

After having carried out the necessary computations, the following results are obtained:



The line of best fit is $Y = mX + C$, where $m = 4.97381078290650$, $C = -23.20729772607083$.

(The vertical lines represent the distance from the line of best fit to the points)

The main limitation of this approach is that in various cases the line of best fit is quite distant from the points. Thus, the approximation in these particular cases will be quite poor.

- The Second Order Approximation

In this method we shall attempt to bend the line of best fit so as to try to accommodate all the given points.

- Build up the table and draw the scatter plot
- Calculate the coefficients m, c for line of best fit
- Calculate the vertical distance from each point to the line of best fit
- If we have 90 numbers, say, in the x -values row, namely 1,2,3,...,90; then sub divide this interval into three intervals, each of equal size. Say first interval contains numbers 1, 2, 3, . . . 30; the second interval contains the numbers 31, 32, 34 60. Similarly for the third interval.
- From each interval choose: 1. the point which is closest to the line, 2. the point which is most distant from the line. In so doing we would end up with three points which are closest to the line in their respective intervals, a_1, a_2, a_3 , and three points which are the most distant from the line 1, 2, 3
- Now we want a quadratic equation of the form $y = ax^2 + bx + c$
- Now we can insert $\alpha_1, \alpha_2, \alpha_3$ in the above equation and find the values of a, b, c which we shall be calling a_1, b_1, c_1 (recall the fact that α_i is a point which contains an x - value and a y -value).
- The same procedure can be done with $\beta_1, \beta_2, \beta_3$ in this case the values we shall obtain are a_2, b_2, c_2 . Now to find the actual values of a, b, c simply compute the following:

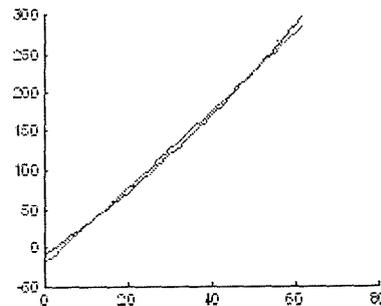
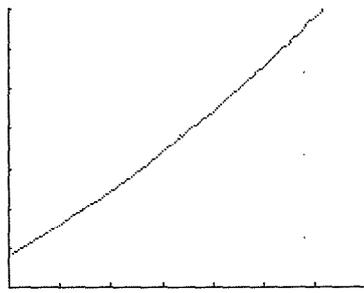
$$a = \frac{a_1 + a_2}{2}; b = \frac{b_1 + b_2}{2}; c = \frac{c_1 + c_2}{2}$$

The equation which can generate the first 300 primes is,

$$y = ax^2 + bx + c$$

where: $a = 0.02070189842422$, $b = 3.68979741019214$, $c = -10.55260457177425$

In the following figures, we can see that the curve now fits the given points much better.



Equations Which can Generate Exact sequences of Prime Numbers

These equations have been created to generate exact sequences of prime numbers. However one must say that although these equations can do this remarkable job, such equations can become very long and rather messy to work with. For example, to generate the first 6 prime numbers we would require a polynomial of order 6. Similarly, to generate the first 100 primes, a polynomial of order 100 would be required. The method used is rather easy and in fact it uses Gaussian elimination.

- Decide for the number of primes that the equation will generate.
- Let the equation be of the form:

$$y = \alpha_n x^n + \alpha_{n-1} x^{n-1} + \dots + \alpha_1 x$$

(where n is the number of primes that the equation is capable of generating)

- To find the coefficients $\alpha_1, \alpha_2, \dots, \alpha_n$ Gaussian elimination can be used.

As an illustration, the equation required to derive the first six prime numbers is:

$$Y = ax^6 + bx^5 + cx^4 + dx^3 + ex^2 + fx$$

where

$$a = -0.020833333$$

$$b = 0.3625$$

$$c = -2.39583$$

$$d = 7.520833333$$

$$e = -11.083$$

$$f = 7.616666667$$