New Formula For the Prime Counting Function

Majid Azimi

Abstract

This paper presents two functions for prime counting function and its inverse function (the function that returns nth prime number as output) with high accuracy and best approximation, which, due to their significant features, are distinguished from other similar functions presented thus far. the presented function for prime counting function is denoted by $\pi_m(x)$ and presented function for nth prime is denoted by $P_m(x)$ in this article.

1 Introduction

The prime number theorem provides a way to approximate the number of primes less than or equal given number n. So one of the most important subject in number theory is the prime counting function that is denoted by $\pi(x)$[1]. Therefore, a lot of functions and theories have been presented in this field.
From Gauss's conjecture[2] in the end of 18th century, which was the first important step in this field ($\pi(x) \simeq \frac{x}{Ln(x)}$), to the Li function and other complex functions have been proposed thus far[3].
Although, many functions have presented for $\pi(x)$, but due to the great computational and structural complexity of the aforementioned and similar functions, the simpler functions with best approximation can be high valuable among others.
Another important subject in prime number theorem is the function that returns nth prime number that is denoted by $P(x)$ in this paper.
Considering the fluctuations and abrupt changes in the graph of the $\pi(x)$ and its inverse function ($P(x)$), discovering a smooth function that always intersects this function with a less computational and structural complexity is very important in several ways, by revealing the count of prime numbers with high accuracy and best possible approximation, it can contribute to better understanding of prime numbers and their distribution.
In this paper, we intend to introduce a smooth function with the above mentioned features that always intersects the graph of the function $\pi(x)$ with less computational and structural complexity compared to similar functions presented thus far. we also intend to introduce a function for $P(x)$ with mentioned features. These functions have a special feature that further distinguishes comparing to its counterparts, which is discussed in the following.
In this paper, the presented prime counting function is denoted by the symbol $\pi_m(x)$, and the function for the nth prime number is denoted by $P_m(x)$.

2 Introducing the Function $\pi_m(x)$

This function is an integral function like Li(x), here is denoted by $\pi_m(x)$. The most important part of this function is a equation denoted by the δ_m with the following definition. δ_m has another use and can be seen in another function that will be introduced in the next section.

$$\delta_m(x) = \gamma\sqrt{x}\ln(x) \tag{1}$$

where γ is the Euler-Mascheroni constant = 0.57721566... (sequence A001620 in the OEIS)

And $\pi_m(x)$ is defined as follows:

$$\pi_m(x) = \int_2^x \frac{dt}{\ln(t + \delta_m(t))} \tag{2}$$

Where

$$\delta_m(x) = \gamma\sqrt{x}\ln(x), \quad \gamma = 0.57721566... \tag{3}$$

So we will have:

$$\pi_m(x) = \int_2^x \frac{dt}{\ln(t + \gamma\sqrt{t}\ln(t))} \tag{4}$$

The below graph compare three functions graphically, $\pi(x)$ (red) , $\pi_m(x)$ (green) and $Li(x)$ (blue).

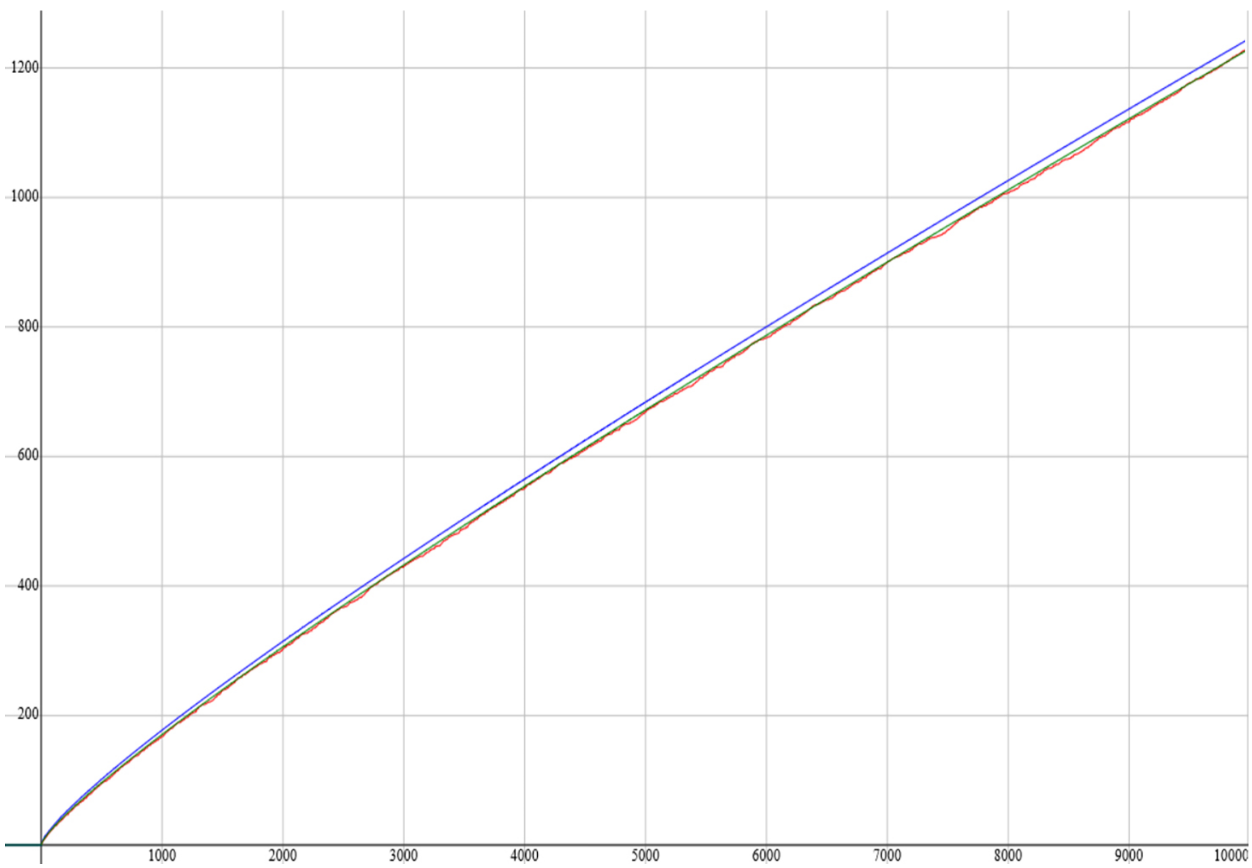

Figure 1: Comparing three Functions, Red: $\pi(x)$, Green: $\pi_m(x)$ and Blue: $Li(x)$

For more accuracy, the below graph shows $\pi(x) - \pi_m(x)$ with red color and $\pi(x) - Li(x)$ with blue color for $(0 < x < 10^7)$, (where the green Line is $y = 0$) :

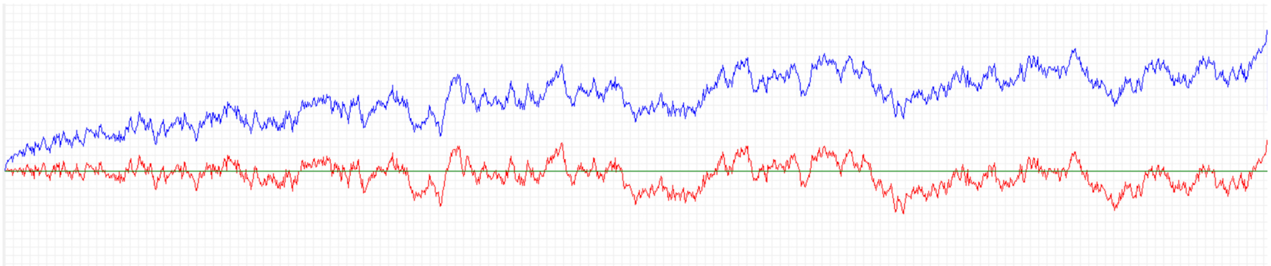

Figure 2: Red Graph: $\pi(x) - \pi_m(x)$ And Blue Graph: $\pi(x) - Li(x)$, for $(0 < x < 10^7)$

As can be seen from the Figure 2, $\pi_m(x)$ function always passes through the function $\pi(x)$ and intersects this function at several points. so it is the best possible approximation for $\pi(x)$.
Note: If any other number between 0.5 and 0.6 is placed instead of γ constant (0.5772...) in the $\pi_m(x)$ equation, the graph still intersects the prime-counting function, but choosing γ constant, in addition to being one of the best coefficients with the highest accuracy, has another reason, which is discussed in the next section.

Below Table shows how the three functions $\pi_m(x)$, $Li(x)$ and $\pi(x)$ compare at powers of 10:

Table 1: Compare three functions $\pi_m(x)$, $Li(x)$ and $\pi(x)$

$\boldsymbol{x}$	$\boldsymbol{\pi}(x)$	$\lfloor \pi_m(x) \rfloor$	$\lfloor \pi_m(x) \rfloor - \boldsymbol{\pi}(x)$	$\lfloor Li(x) \rfloor - \pi(x)$
10^1	4	4	0	2
10^2	25	26	+1	5
10^3	168	170	+2	10
10^4	1,229	1,229	0	17
10^5	9,592	9,590	-2	38
10^6	78,498	78,526	+28	130
10^7	664,579	664,653	+74	339
10^8	5,761,455	5,761,492	+38	754
10^9	50,847,534	50,847,255	-279	1,701

3 Introducing the Function $P_m(x)$

in this paper, the function that return nth Prime number as output (for $n \geq 1$ so that n is integer) is denoted by $P(x)$. For example: $P(1) = 2, P(2) = 3, P(3) = 5, P(4) = 7, P(5) = 11, ...$
Now, we assume the $P(x)$ function is a *differentiable* and *continuous* function so we could integrate.
Although it is debatable whether a *continuous function* that returns nth prime number could exist or not, we assume there is a continuous function that return nth prime number, because we want to integrate of that in the next equation.
Now, According to the continuity definition, $P(x)$ must return the value as output for none-integer numbers

and does not have any abrupt changes in value. So, we can make this condition(differentiable) by drawing a straight line from $y = P(x)$ to $y = P(x + 1)$ (where x is integer number) on the graph. as shown in the figure below:

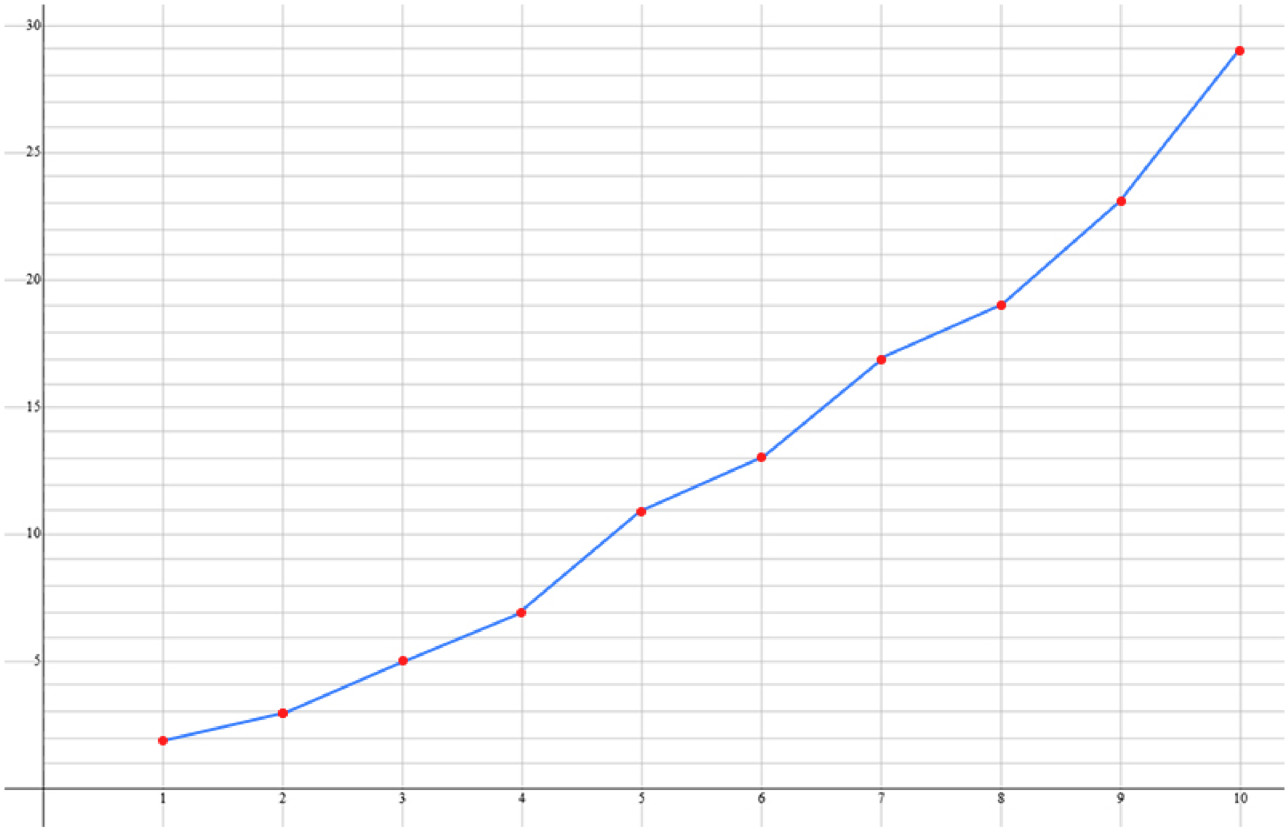

Figure 3: Graph of Smooth Function for $P(x)$

With this assumption, $P(x)$ is calculated using the following formula for non-integer numbers:

$P(x) = P([x]) + (x - [x]) \cdot gap(x) \ \ \Rightarrow$ where $gap(x) = P_{x+1} - P_x$

In this case, Like $\pi_m(x)$ function which always intersects the prime-counting function on the graph, a function can be defined that has the same feature for the inverse function, and thus always intersects $P(x)$. This function is denoted by $P_m(x)$ in this paper with below definition:

$$P_m(x) = \int_1^x \ln(P(t) + \delta_m(P(t))) \cdot dt \tag{5}$$

Where

$$\delta_m(x) = \gamma\sqrt{x}\ln(x), \quad \gamma = 0.57721566... \tag{6}$$

And $P(x)$ is xth Prime Number Function
So we will have:

$$P_m(x) = \int_1^x \ln(P(t) + \gamma\sqrt{P(t)}\ln(P(t))) \cdot dt \tag{7}$$

The below graph compare tow function $P(x)$ and $P_m(x)$ by $P(x) - P_m(x)$ function:

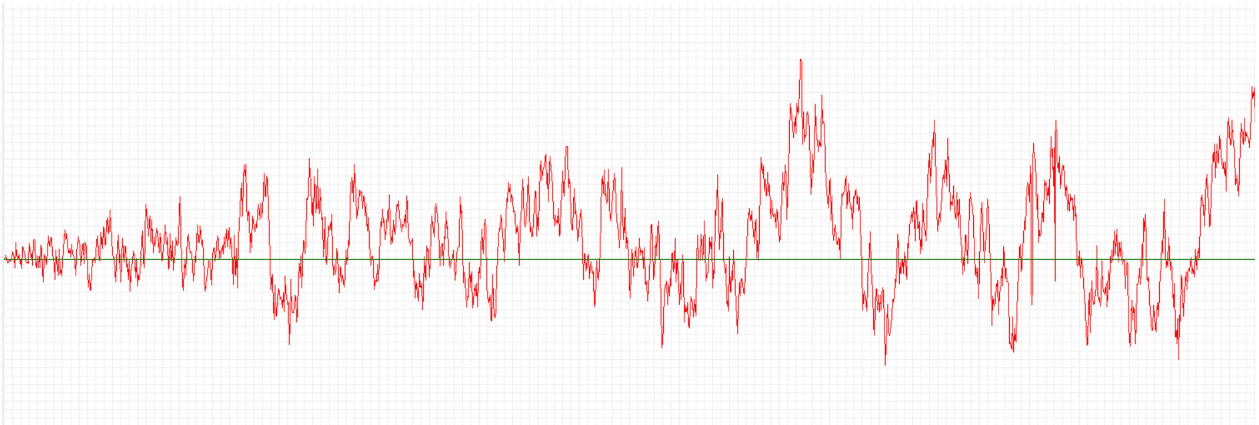

Figure 4: Graph of $P(x) - P_m(x)$, for $0 < x < 10^7$

As can be seen from the Figure 4, $P_m(x)$ function always passes through the function $P(x)$ and intersects this function at several points.
Though $P_m(x)$ has a high complex structure (due to using integral of $P_(x)$ in the structure) it may seem unimportant, but it does matter in a way.
As mentioned before, the δ_m function is used in the $P_m(x)$ function as well as its inverse function (π_m) and as can be seen there is a significant mathematical relationship between $\pi_m(x)$ and its inverse function: $P_m(x)$

Note: if any other number between 0.5 and 0.6 is placed instead of γ constant (0.5772...), the graph still intersects the $P(x)$ function (Like the $\pi_m(x)$), but choosing γ constant, in addition to being one of the best coefficients with the highest accuracy, has another reason, which is discussed in the next section.

4 Reasons for choosing γ as coefficient in introduced functions

In addition to the reason mentioned before, if we want to assume that we have to choose same coefficient with the highest accuracy for both functions, a number close to the γ is the best possible candidate.
This means that if we consider this coefficient to be slightly larger than γ (e.g. $\gamma + 0.05$), the function $\pi_m(x) - \pi(x)$ will be plotted a little above the $y = 0$ line .And if, conversely, we choose this coefficient slightly smaller than γ (e.g. $\gamma - 0.05$), the function $P_m(x) - P(x)$ will be plotted a little below the $y = 0$ line.
In short, the γ Constant is a weighted number balancing these two functions. Obviously, we can consider two different coefficient for each of these two functions which may seem a little more accurate to each of them individually, but this differences in accuracy is very fiddling which can be ignored. therefore, considering this issue and principle of beauty in mathematics, it seems that choosing γ as coefficient is a more correct and attractive choice.

References

[1] Bach, Eric;Shallit, Jeffery (1996) *Algorithmic Number Theory*. MIT Press. volume 1 page 234 section 8.8. ISBN 0-262-02405-5.

[2] Dominic Klyve (2019) *The Origin of the Prime Number Theorem*, MAA Convergence

[3] Tadej Kotnik. *The prime-counting function and its analytic approximations*.Springer Science + Bussiness Media B.V .2007