波利尼亚克猜想：数学表达式的一个推导

胡坚

San Francisco, U. S. A., July 30, 2024
Email: jian.hu6624@gmail.com

摘要

本文运用分段式筛法（segmented Eratosthenes' sieve）推导出最先由法国数学家波利尼亚克（Alphonse de Polignac）于 1849 年提出的猜想，即任意相邻的两个素数之差等于一个偶数，

$$p_{k+1} = p_k + 2h. \quad (h=1, 2, 3, \cdots)$$

可用数学归纳法证明，对于任意k值，上述公式均成立。。

1. 序言

波利尼亚克猜想（Polignac's conjecture）最早由法国数学家波利尼亚克（Alphonse de Polignac）于1849年发表的两篇论文中所提出：一篇是他在该年10月15日提交给法兰西科学院的研究报告[1]，而另一篇是同年底他发表在法国《纳维尔数学年刊》（Nouvelles Annales de Mathématiques）上的论文[2]，后者基本上是前者的一个较短的版本。在这些论文（或报告）的最后一节，作者写道："第1定理：存在无穷多组两个相邻素数对，其差等于一个偶数。"虽然作者称之为"定理"，但其论文本身并无推导或是证明，故后世数学家一般都把波利尼亚克的结果看作是猜想而非经过确证的定理。本文中，作者试图基于筛法理论，直接推导出相邻素数与任意偶数之间的关系式。在下面的一节，我们首先描述分段式筛法（segmented Eratosthenes' sieve），由此产生素数按分段长度的分布图；接下来根据该分布图，推导出相邻素数与任意偶数之间的公式（也即波利尼亚克猜想的数学描述）并用数学归纳法予以证实。

2. 分段式筛法和不寻常素数分布图

在分段式筛法中，自然数被分割成不同大小的数值间隙（intervals）[3]：$S_1(4,8)$, $S_2(9,24)$, $S_3(25,48)$, … , $S_k(p_k^2\ , p_{k+1}^2 - 1)$（$k \geq 1$）。为求得每个简隙内的素数，须剔除不大于$\sqrt{p_k^2}$ (即p_k) 的所有素数$\{p_1, \dots , p_k\}$的倍数。另外定义“种子”区间$S_0(1,3)$，因为该区间没有任何合数，故过筛后得到两个素数2和3。接着在区间S_1、S_2、S_3、…分别剔除素数{2}、{2,3}、{2,3,5}、…的倍数，直至无任何合数存在。值得指出的是，每区间下限值的平方根正好等于用来筛选该区间的素数组的最后那个素数，如对于$S_2(9,24)$是素数3，而在$S_3(25,48)$则是素数5，等等。若定义$p_{k,last}$为素数组的最后那个素数，则有下式：

$$p_k = p_{k,last}. \ (k \geq 1) \tag{1}$$

式中p_k是每区间的下限值平方根。

表1示出了采用分段式筛法获取素数的部分结果（因篇幅有限，仅列出前面5个区间的结果）。p_k^2 和 p_{k+1}^2 分别为区间S_k的上、下限值；l_k是区间S_k的长度，定义为 $l_k := |S_k| = p_{k+1}^2 - p_k^2$ [6]。

Table 1: 在前5区间(S_1~S_5)的筛选结果。来自初始区间 (S_0) 的结果也一并列出。

$\boldsymbol{k}$	$\boldsymbol{p_{k,last}}$	$\boldsymbol{p_{k+1}^2}$	$\boldsymbol{l_k}$	**每个区间$\boldsymbol{S_k}$产生的素数**
(初始区间)	2	4	3	S_0: 2{2,3}
1	2	9	5	S_1: 2{5,7}
2	3	25	16	S_2: 5{11,13,…,23}
3	5	49	24	S_3: 6{29,31,…,47}
4	7	121	72	S_4: 15{53,59,…,113}
5	11	169	48	S_5: 9{127,131,…167}
…				

现在可绘出素数$p_{k,last}\ vs.\ l_k$图，或$p_k\ vs.\ l_k$图（根据式（1），$p_{k,last}$和p_k相同），如图1所示。

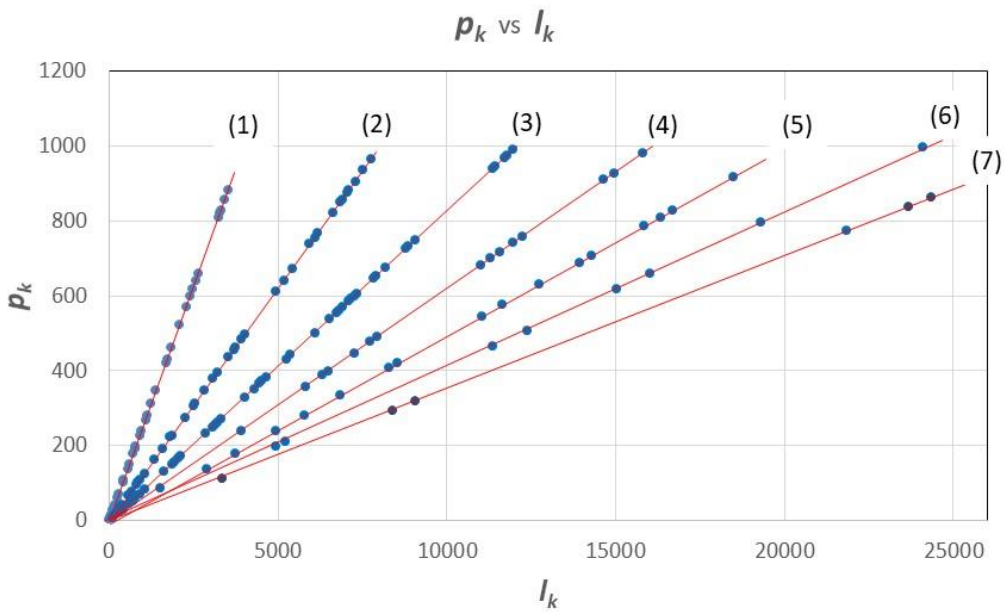

Fig.1. $p_k\ vs.\ l_k$. 为简明见，只使用了起始的167素数，即p_k={3, 5, …, 997}。注意素数2因为不能落在任何曲线，故没有包括。

令人惊奇地是，所有素数并不是落在同一曲线上，而是分布在不同曲线上；每个素数也只能落在其中一个曲线上（不能同时落在两曲线上）。另外，所有曲线能被一简单的线性回归模型描述，且其斜率和截距呈有规律地变化：

$$
\begin{aligned}
\text{直线 (1)}: p_{k,1} &= l_k/4 - 1;\\
(2): p_{k,2} &= l_k/8 - 2;\\
(3): p_{k,3} &= l_k/12 - 3;\\
(4): p_{k,4} &= l_k/16 - 4;\\
&\ldots\ldots
\end{aligned}
$$

遵从如下一般线性方程：

$$p_{k,h} = l_k/4h - h,\ (h = 1,2,3, \ldots) \qquad (2)$$

式中h是一正整数，不随k变化。很显然图1中的那些直线能被h唯一地确定：当h增加时，相邻直线的斜率按$1/h$因子减少；其截距也随之减少h单位。图1中显示出的那种奇特的素数分布其实也是因为受h的影响所致：即每条直线代表了具有同一h值的所有素数，例如$h = 1$的所有素数都落在直线（1）上，所有$h = 2$的素数落在直线（2）上，…等等。而当一新的直线出现时，意味着必须要有一由不同素数所共有的相对应的新的h值。

3. 相邻素数与偶数的相关性

基于式（2），我们现在可以推导出相邻素数与偶数之间的相关性。根据定义，区间长度 $l_k = p_{k+1}^2 - p_k^2$, 将此式代入式（2）并消掉l_k，可得

$$p_{k+1} = p_k + 2h.\ \ (h = 1,2,3, \dots) \tag{3}$$

此式正是波利尼亚克猜想中所描述的“两个相邻素数对，其差等于一个偶数”！从上面的分析推导过程，我们可清楚地看到区间长度（l_k）从中起了非常关键的作用，因为正是从$p_k\ vs.\ l_k$图（图1），我们才发现了不寻常的素数分布并推导出式（2）和式（3），所有这些无疑得益于分段式筛法，而传统的筛法（例如波利尼亚克所使用的那种）是难以做到的。

式（3）对于任意素数的可适用性可通过数学归纳法进行证实。考虑素数系列 $\{p_1, p_2, p_3, \dots p_n\}$（$p_1 = 2$）。如前所述，素数2不能归于任何$p_k \sim l_k$直线，故可忽略，而从 $p_2 = 3$开始。假如$n = 2$, 则从式（3）, $p_3 - p_2 = 5 - 3 = 2$. 因此对于$n = 2$，式（3）成立。现在又假设当 $n = k$时， $p_{k+1} - p_k = 2h\ \ (h = 1,2,3, \dots)$也成立，则从式（2）， 有

$$p_{k+1} = l_{k+1}/4h - h,\ \text{或}$$

$$l_{k+1} = 4hp_{k+1} + 4h^2 = 4hp_k + 12h^2. \tag{4}$$

另外根据定义, $l_{k+1} = p_{k+2}^2 - p_{k+1}^2$, 因此,

$$l_{k+1} = p_{k+2}^2 - p_{k+1}^2 = p_{k+2}^2 - (p_k^2 + 4hp_k + 4h^2) \tag{5}$$

从式(4) and (5)消掉l_{k+1}, 则有

$$p_{k+2}^2 = p_k^2 + 8hp_k + 16h^2 = (p_k + 4h)^2,$$

进而得到 $p_{k+2} - p_k = 4h$, 或 $p_{k+2} - p_{k+1} = 2h$. 所以式（3）对于$n = k + 1$也成立。这就证明了式（3）对于任意素数均成立。

4. 结论

(1) 采用分段式筛法发现素数p_k（>2）与区间长度l_k之间存在着不寻常的相关性，可用下面线性方程准确地描述: (式（2）)

$$p_{k,h} = l_k/4h - h.\ (h = 1,2,3,\ldots)$$

(2) 基于式（2），推导出如下描述波利尼亚克猜想的数学公式：（式（3））

$$p_{k+1} = p_k + 2h.\ (h = 1,2,3,\ldots)$$

其中p_k、p_{k+1}是两个相邻且均大于2的素数。数学归纳法证明，式（3）对于任意素数均适用。

参考文献

[1] De Polignac, M. Alphonse, Recherches nouvelles sur les no'nbres premiers, *Comptes rendus hebdomadaires des séances de l'Académie des Sciences*, October 15, 1849, 397-401.

[2] De Polignac, M. Alphonse, Six propositions arithmologiques déduites de crible d'Ératosthène. *Nouv. Ann. Math.*, 1849, **8**: 428-429. Note on page 427, the author had specifically referred to his earlier report.

[3] Tunstrom, Kolbjorn, The origin of the logarithmic integral in the prime number theorem. 2013, **arXiv:1311.1093:** 1-31.