漫谈 Bertrand 假设

李润博 (Runbo Li)
2025/11/21
南开中学

素数起源

一个大于 1 的正整数如果只有两个因子（1 和它本身），则它是一个素数。

素数起源

一个大于 1 的正整数如果只有两个因子（1 和它本身），则它是一个素数。

素数的概念可以追溯到几千年前的古代文明，古希腊的 Euclid 通过一个巧妙的反证法证明了素数有无穷多个：假设素数只有有限多个，则必然存在一个最大素数。把从 2 到这个数的所有素数乘起来再 $+1$，得到的这个数不能被所有已知的素数整除，因此它必然是一个更大的素数，得到矛盾。

素数起源

一个大于 1 的正整数如果只有两个因子（1 和它本身），则它是一个素数。

素数的概念可以追溯到几千年前的古代文明，古希腊的 Euclid 通过一个巧妙的反证法证明了素数有无穷多个：假设素数只有有限多个，则必然存在一个最大素数。把从 2 到这个数的所有素数乘起来再 $+1$，得到的这个数不能被所有已知的素数整除，因此它必然是一个更大的素数，得到矛盾。

由于素数的独特性，它在正整数中的分布一直是数学家们关注的永恒的问题。

下面我们默认 p 表示素数，$\pi(x)$ 表示不超过 x 的素数个数，$\log$ 表示自然对数，ε 表示充分小的正数。

Bertrand 假设

1845 年，法国数学家 Bertrand 提出了一个关于素数的猜想：是否对所有正整数 x，区间 $[x, 2x]$ 之间都存在素数？他本人对于不超过 3000000 的正整数进行了验证。

Bertrand 假设

1845 年，法国数学家 Bertrand 提出了一个关于素数的猜想：是否对所有正整数 x，区间 $[x, 2x]$ 之间都存在素数？他本人对于不超过 3000000 的正整数进行了验证。

1852 年，俄国数学家 Chebyshev 给出了这个猜想的第一个完整证明，此后这个命题被称为 **Bertrand 假设**或 **Bertrand-Chebyshev 定理**。

Bertrand 假设

1845 年，法国数学家 Bertrand 提出了一个关于素数的猜想：是否对所有正整数 x，区间 $[x, 2x]$ 之间都存在素数？他本人对于不超过 3000000 的正整数进行了验证。

1852 年，俄国数学家 Chebyshev 给出了这个猜想的第一个完整证明，此后这个命题被称为 **Bertrand 假设**或 **Bertrand-Chebyshev 定理**。

此后，很多数学家对这一命题给出了许多不同的证明，包括印度数学家 Ramanujan 等人。其中最有名的当属匈牙利数学家 Paul Erdős 于 1932 年通过估计中心二项式系数 $\binom{2n}{n}$ 给出的初等证明。Erdős 的证明可以在 *Number Theory: Concepts and Problems* 中找到。下面我们将介绍这个证明。

Erdős 对 Bertrand 假设的证明

设 P_n 表示 n 到 $2n$ 之间的所有素数的乘积（若区间内没有素数则为 1）。我们将证明一个更强的定理：对所有 $n > 225$ 成立，有 $P_n > 4^{\frac{n}{3}}(2n)^{-\sqrt{\frac{n}{2}}}$。通过验证所有 $n \leq 225$，我们知道 Bertrand 假设是这个定理的推论。在后面的证明中我们默认 $n > 225$ 。

设 $A = \binom{2n}{n}$，显然 A 的所有素因子在 1 和 $2n$ 之间。不难发现当 $n < p \leq 2n$ 时，$v_p(A) = 1$，因此

$$A = P_n \prod_{p \leq n} p^{v_p(A)}.$$

Erdős 对 Bertrand 假设的证明

由于二项式系数的定义，我们知道

$$\sum_{k=0}^{2n} \binom{2n}{k} = 4^n.$$

因为中心二项式系数是求和里最大的一项，我们有估计

$$A \geq \frac{4^n}{2n+1} > \frac{4^{n-1}}{2n}.$$

因此要证定理成立，只需证明

$$\prod_{p \leq n} p^{v_p(A)} < 4^{\frac{2n}{3}-1} (2n)^{\sqrt{\frac{n}{2}}-1}.$$

Erdős 对 Bertrand 假设的证明

在证明定理之前，我们先介绍几个引理。

引理 1（Legendre 公式）. 设 $v_p(n)$ 表示 n 的含 p 量（等于最大的正整数 r，满足 $p^r \mid n$）。则对所有素数 p 和所有正整数 n，有

$$v_p(n!) = \lfloor \frac{n}{p} \rfloor + \lfloor \frac{n}{p^2} \rfloor + \lfloor \frac{n}{p^3} \rfloor + \cdots$$

注意到这个无穷求和实际上只有有限多个非零项，因为一定存在 k 使得 $p^k > n$。

Erdős 对 Bertrand 假设的证明

引理 1 的证明. 我们有

$$v_p(n!) = v_p(1 \times 2 \times \cdots \times n) = v_p(1) + v_p(2) + \cdots + v_p(n).$$

在 $1, 2, \ldots, n$ 这些数中，有 $\lfloor \frac{n}{p} \rfloor$ 个 p 的倍数，$\lfloor \frac{n}{p^2} \rfloor$ 个 p^2 的倍数，以此类推。其中是 p 的倍数但不是 p^2 的倍数在求和中的贡献是 1，其中是 p^2 的倍数但不是 p^3 的倍数在求和中的贡献是 2，以此类推。因此，

$$v_p(n!) = \left(\lfloor \frac{n}{p} \rfloor - \lfloor \frac{n}{p^2} \rfloor \right) + 2 \left(\lfloor \frac{n}{p^2} \rfloor - \lfloor \frac{n}{p^3} \rfloor \right) + \cdots$$

展开求和，引理 1 证毕。

Erdős 对 Bertrand 假设的证明

引理 2. 设整数 $n \geq k \geq 0$，则 $p^{v_p(\binom{n}{k})} \leq n$。也就是说，所有整除 $\binom{n}{k}$ 的素数幂不超过 n。

引理 2 的证明. 引理 1 给出

$$v_p(\binom{n}{k}) = v_p(n!) - v_p(k!) - v_p((n-k)!) = \sum_{j\geq 1}\left(\lfloor\frac{n}{p^j}\rfloor - \lfloor\frac{k}{p^j}\rfloor - \lfloor\frac{n-k}{p^j}\rfloor\right).$$

由基本不等式

$$\lfloor x+y\rfloor - \lfloor x\rfloor - \lfloor y\rfloor \in \{0,1\}$$

对实数 x, y，求和中的每一项只能是 0 或 1。当 $p^j > n$ 时，这一项里的三个下取整均为 0。而前面的求和中非零项至多有 $\lfloor \log_p(n)\rfloor$ 个，因此 $v_p(\binom{n}{k}) \leq \lfloor \log_p(n)\rfloor$，引理 2 证毕。

Erdős 对 Bertrand 假设的证明

引理 3（Erdős）. 设 $n \geq 2$，则不超过 n 的所有素数乘积小于 4^{n-1}。

引理 3 的证明. 用数学归纳法。$n = 2$ 时显然成立，下面假设结果直到 $n-1$ 都成立，若 n 是偶数，则新加的一项不可能是素数，因此我们只需考虑 n 是奇数的情况。设 $n = 2k+1$。因为

$$\binom{2k+1}{k} = \frac{(k+2)(k+3)\cdots(2k+1)}{k!}$$

含有所有 $k+2 \leq p \leq n$ 的素数为因子，所以

$$\prod_{p \leq n} p \leq \prod_{p \leq k+1} p \binom{2k+1}{k}.$$

Erdős 对 Bertrand 假设的证明

用组合数的定义，我们知道

$$\binom{2k+1}{k} < 4^k.$$

根据归纳假设 $\prod_{p \le k+1} p < 4^k$ 和估计 $\binom{2k+1}{k} < 4^k$，因此

$$\prod_{p \le n} p \le 4^k \cdot 4^k = 4^{n-1}.$$

引理 3 证毕。

Erdős 对 Bertrand 假设的证明

下面我们将证明定理。根据引理 2，每个 $p^{v_p(A)}$ 不超过 $2n$。而引理 1 给出 $v_p(A) \le 1$ 对 $p > \sqrt{2n}$ 成立。对 $p \in (\frac{2n}{3}, n]$，有 $v_p((2n)!) = 2$ 和 $v_p(n!) = 1$，因此对 $p \in (\frac{2n}{3}, n]$，有 $v_p(A) = 0$。综合起来有

$$\prod_{p \le n} p^{v_p(A)} \le \prod_{p \le \sqrt{2n}} (2n) \prod_{\sqrt{2n} < p \le \frac{2n}{3}} p.$$

Erdős 对 Bertrand 假设的证明

现在设 $n \geq 125$，$k = \lfloor \sqrt{2n} \rfloor \geq 15$。因为 $1, 9, 15$ 和 $4, \ldots, 2\lfloor \frac{k}{2} \rfloor$ 都不是素数，我们有

$$\pi(k) \leq k - (2 + \lfloor \frac{k}{2} \rfloor) < \frac{k}{2} - 1 \leq \sqrt{\frac{n}{2}} - 1.$$

结合引理 3 最后给出

$$\prod_{p \leq n} p^{v_p(A)} < 4^{\frac{2n}{3} - 1} (2n)^{\sqrt{\frac{n}{2}} - 1}.$$

定理证毕。

Bertrand 假设的改进

很自然地，我们思考能否将区间 $[x, 2x]$（长度为 x）进一步缩小，使得其中仍然有素数?

Bertrand 假设的改进

很自然地，我们思考能否将区间 $[x, 2x]$（长度为 x）进一步缩小，使得其中仍然有素数?

最简单的想法是考虑区间 $[2y, 3y]$（$y = \frac{1}{2}x$ 时等价于 $[x, \frac{3x}{2}]$），此时区间长度为 $\frac{x}{2}$，比原先短了一半。通过与 Erdős 的证明中类似的技术，我们仍然可以证明其中存在素数。

Bertrand 假设的改进

很自然地，我们思考能否将区间 $[x, 2x]$（长度为 x）进一步缩小，使得其中仍然有素数?

最简单的想法是考虑区间 $[2y, 3y]$（$y = \frac{1}{2}x$ 时等价于 $[x, 1.5x]$），此时区间长度为 $0.5x$，比原先短了一半。通过与 Erdős 的证明中类似的技术，我们仍然可以证明其中存在素数。

实际上，Nagura 在 1953 年就已经证明了，对于任意正整数 $x > 25$，区间 $[x, x + \frac{1}{5}x]$ 中都存在素数。

Bertrand 假设的改进

很自然地，我们思考能否将区间 $[x,2x]$（长度为 x）进一步缩小，使得其中仍然有素数?

最简单的想法是考虑区间 $[2y,3y]$（$y=\frac{1}{2}x$ 时等价于 $[x,1.5x]$），此时区间长度为 $0.5x$，比原先短了一半。通过与 Erdős 的证明中类似的技术，我们仍然可以证明其中存在素数。

实际上，Nagura 在 1953 年就已经证明了，对于任意正整数 $x>25$，区间 $[x,x+\frac{1}{5}x]$ 中都存在素数。

此后，由于区间长度相对于 x 会非常小，我们不再考虑命题"对所有正整数成立"，而考虑命题在"对某个正整数以上的正整数都成立"（实效），或"对充分大的正整数都成立"（非实效）。

Bertrand 假设的改进 (实效)

前面提到，Nagura 在 1953 年证明了区间 $[x, x+\frac{1}{5}x]$ 中存在素数对任意正整数 $x > 25$ 成立。我们能否在对 x 的下界不做过多要求的情况下把 $\frac{1}{5}$ 改进为更小?

Bertrand 假设的改进 (实效)

前面提到，Nagura 在 1953 年证明了区间 $[x, x + \frac{1}{5}x]$ 中存在素数对任意正整数 $x > 25$ 成立。我们能否在对 x 的下界不做过多要求的情况下把 $\frac{1}{5}$ 改进为更小?

2015 年，Balliet 用初等方法证明了：区间 $[x, x + \frac{1}{n}x]$ 中存在素数对任意正整数 $n \le 8$ 及 $x > n^2$ 成立。使用非初等的方法，他进一步把范围扩大到 $n \le 519$。

如果我们允许 $x > 2010760$，则 Schoenfeld 证明了 $n \le 16597$。

素数定理

1896 年，两位数学家 Hadamard 和 de la Vallée Poussin 各自独立证明了著名的素数定理：

$$\pi(x) \sim \frac{x}{\log x}.$$

这个定理直接表明对于充分大的 x，区间 $[x, x + \varepsilon x]$ 中存在素数。

Bertrand 假设的改进 (实效)

上面提到的区间长度最短为 εx，但从量级上看仍然和 x 一样。我们能否将这一长度缩小为比 x 的量级更小?

Bertrand 假设的改进 (实效)

上面提到的区间长度最短为 εx，但从量级上看仍然和 x 一样。我们能否将这一长度缩小为比 x 的量级更小?

在 1998 年，Dusart 证明了对任意 $x > 3275$，区间 $[x, x + \frac{1}{2\log^2 x}x]$ 中存在素数。

Bertrand 假设的改进 (实效)

上面提到的区间长度最短为 εx，但从量级上看仍然和 x 一样。我们能否将这一长度缩小为比 x 的量级更小?

在 1998 年，Dusart 证明了对任意 $x > 3275$，区间 $[x, x + \frac{1}{2\log^2 x}x]$ 中存在素数。

在 2010 年，他又证明了对任意 $x > 396738$，区间 $[x, x + \frac{1}{25\log^2 x}x]$ 中存在素数。

Bertrand 假设的改进 (实效)

上面提到的区间长度最短为 εx，但从量级上看仍然和 x 一样。我们能否将这一长度缩小为比 x 的量级更小?

在 1998 年，Dusart 证明了对任意 $x > 3275$，区间 $[x, x + \frac{1}{2\log^2 x}x]$ 中存在素数。

在 2010 年，他又证明了对任意 $x > 396738$，区间 $[x, x + \frac{1}{25\log^2 x}x]$ 中存在素数。

在 2016 年，他又又证明了对任意 $x > 468991632$，区间 $[x, x + \frac{1}{5000\log^2 x}x]$ 中存在素数。

Bertrand 假设的改进 (实效)

上面提到的区间长度最短为 εx，但从量级上看仍然和 x 一样。我们能否将这一长度缩小为比 x 的量级更小?

在 1998 年，Dusart 证明了对任意 $x > 3275$，区间 $[x, x + \frac{1}{2\log^2 x}x]$ 中存在素数。

在 2010 年，他又证明了对任意 $x > 396738$，区间 $[x, x + \frac{1}{25\log^2 x}x]$ 中存在素数。

在 2016 年，他又又证明了对任意 $x > 468991632$，区间 $[x, x + \frac{1}{5000\log^2 x}x]$ 中存在素数。

同年，他又又又证明了对任意 $x > 89693$，区间 $[x, x + \frac{1}{\log^3 x}x]$ 中存在素数。

Bertrand 假设的改进 (非实效)

如果我们不去考虑 x 的下界，而只关心当 x 充分大时的情况，则区间长度会比实效版本小很多。

Bertrand 假设的改进 (非实效)

如果我们不去考虑 x 的下界，而只关心当 x 充分大时的情况，则区间长度会比实效版本小很多。

早在 1930 年，Hoheisel 就证明了对任意充分大的 x，区间 $[x, x + x^{32999/33000}]$ 中都存在素数。这比 Dusart 等人得到的实效结果好不少。

Bertrand 假设的改进 (非实效)

如果我们不去考虑 x 的下界，而只关心当 x 充分大时的情况，则区间长度会比实效版本小很多。

早在 1930 年，Hoheisel 就证明了对任意充分大的 x，区间 $[x, x + x^{32999/33000}]$ 中都存在素数。这比 Dusart 等人得到的实效结果好不少。

此后，许多数学家致力于缩小区间 $[x, x + x^{\theta}]$ 中 θ 的大小，目前最好的发表结果是由三位数学家 Baker、Harman 和 Pintz 在 2001 年得到的 $\theta \geq 0.525$。李润博在一篇预印本中得到了 $\theta \geq 0.52$。

Bertrand 假设的改进 (实效)

上面提到的实效版本的区间长度最短已经达到了 $x\log^{-3}x$，量级上小于 x (这意味着当 $x\to\infty$ 时区间长度远远小于任意 cx)，但仍然大于 $x^{1-\varepsilon}$。

Bertrand 假设的改进 (实效)

上面的区间长度最短已经达到了 $x\log^{-3}x$，量级上小于 x（这意味着当 $x\to\infty$ 时区间长度远远小于任意 cx），但仍然大于 $x^{1-\varepsilon}$。

2014 年，Dudek 首次证明了：对任意 $x>\exp(\exp(33.217))$，区间 $[x,x+x^{2/3}]$ 中存在素数。

这是一个惊人的结果。

Bertrand 假设的改进 (实效)

上面的区间长度最短已经达到了 $x\log^{-3}x$，量级上小于 x（这意味着当 $x\to\infty$ 时区间长度远远小于任意 cx），但仍然大于 $x^{1-\varepsilon}$。

2014 年，Dudek 首次证明了：对任意 $x>\exp(\exp(33.217))$，区间 $[x,x+x^{2/3}]$ 中存在素数。

这是一个惊人的结果。

此后，x 的下界被 Mattner、Cully-Hugill 等人先后降到 $x>\exp(\exp(32.537))$，但区间长度 $x^{2/3}$ 仍然没有被改进。前面提到的三位数学家 Baker、Harman 和 Pintz 认为他们的 0.525 也可以做出实效版本。

猜想结果

著名的 Riemann 猜想能够推出，对任意充分大的 x，区间 $[x, x + x^{1/2}\log x]$ 中都存在素数。

猜想结果

著名的 Riemann 猜想能够推出，对任意充分大的 x，区间 $[x, x + x^{1/2}\log x]$ 中都存在素数。

稍弱一些的 Lindelöf 猜想（Riemann 猜想的推论）能够推出，对任意充分大的 x，区间 $[x, x + x^{1/2+\varepsilon}]$ 中都存在素数。

猜想结果

著名的 Riemann 猜想能够推出，对任意充分大的 x，区间 $[x, x + x^{1/2}\log x]$ 中都存在素数。

稍弱一些的 Lindelöf 猜想（Riemann 猜想的推论）能够推出，对任意充分大的 x，区间 $[x, x + x^{1/2+\varepsilon}]$ 中都存在素数。

Legendre 猜想说两个相邻平方数之间必有素数，这等价于对任意 $x > 1$，区间 $[x, x + x^{1/2}]$ 中都存在素数，由此可见它强于黎曼猜想。

目前，人们距离这些猜想还相差甚远。

猜想结果

著名的 Riemann 猜想能够推出，对任意充分大的 x，区间 $[x, x+x^{1/2}\log x]$ 中都存在素数。

稍弱一些的 Lindelöf 猜想（Riemann 猜想的推论）能够推出，对任意充分大的 x，区间 $[x, x+x^{1/2+\varepsilon}]$ 中都存在素数。

Legendre 猜想说两个相邻平方数之间必有素数，这等价于对任意 $x > 1$，区间 $[x, x+x^{1/2}]$ 中都存在素数，由此可见它甚至强于黎曼猜想。

目前，人们距离这些猜想还相差甚远。

2004 年，Friedlander 和 Iwaniec 在假设一个有关 L-函数例外特征存在的猜想下，证明了区间 $[x, x+x^{0.4937}]$ 中存在素数。李润博在一篇预印本中把它改进到 $[x, x+x^{0.4923}]$。

几乎所有区间内的素数

如果我们不要求区间 $[x, x + x^{\theta+\varepsilon}]$ 对所有充分大的 x 都存在素数，而只要求在 $x \in [X, 2X]$ 的情况下，不存在素数的这种区间占比为 0，则可以得到更好的结果。

几乎所有区间内的素数

如果我们不要求区间 $[x, x+x^{\theta+\varepsilon}]$ 对所有充分大的 x 都存在素数，而只要求在 $x \in [X, 2X]$ 的情况下，不存在素数的这种区间占比为 0，则可以得到更好的结果。

1943 年，Selberg 最早证明了 θ 可以取 $\frac{19}{77}$。

目前最好的发表结果是贾朝华在 1996 年得到的 $\frac{1}{20}$。

李润博在三篇预印本中先后得到了 $\frac{1}{21.5}$，$\frac{1}{22}$ 和 $\frac{1}{24}$。

素数的有界间隔

2013 年，张益唐证明了存在无穷多个素数对相差不超过 70000000，即存在无穷多个长度为 70000000 的包含至少 2 个素数的整数区间。这一结果被 Maynard，Polymath 团队等人先后改进到长度为 246。

存在无穷多个包含至少 3 个素数或更多素数的整数区间长度的最好结果是 Stadlmann 在 2023 年得到的。

素数的大间隔

设 $G(x)$ 表示不超过 x 的所有相邻素数对中间隔最长的长度。目前 $G(x)$ 的最好下界是 2017 年由 Ford、Green、Konyagin、Maynard 和 Tao 证明的

$$G(x) \gg \frac{\log x \log\log x \log\log\log\log x}{\log\log\log x}.$$

Erdős 对这个问题的猜想下界是

$$G(x) \gg (\log x)^{1+\varepsilon}.$$

他对这个猜想悬赏 1 万美金。

素数的大间隔

$G(x)$ 下界的证明可以归结到另一个类似问题的证明（图中的问题是弱化版本）：

对正整数 n，用 P_n 表示全体不大于 n 的质数的乘积. 求证：存在常数 $c>0$，使得对任意充分大的整数 n，若将全体不大于 P_n 且与 P_n 互质的数排成一列，则存在相邻两个数的差不小于 $c\cdot\dfrac{n\cdot\ln n}{(\ln\ln n)^2}$，并考虑上述量级能否加强为 $c\cdot\dfrac{n\cdot\ln n\cdot\ln\ln\ln n}{(\ln\ln n)^2}$.（天津实验中学滨海学校 孙牧聪 供题）

这是 2023 年 9 月数之谜平台比赛“谜之竞赛”的第 7 题，由孙牧聪提供。图中的两个问题分别对应 Erdős 和 Rankin 在 1935 年和 1938 年得到的结果。

学习数学的一点建议

- 学自己喜欢的数学
- 多思考、多见识、多和人交流
- 学会自行查找资料
- 中文平台：数之谜（微信小程序）
- 英文平台：MathOverflow、MathStackExchange

数论入门推荐书籍

- *Number Theory: Concepts and Problems* by Titu Andreescu, Gabriel Dospinescu and Oleg Mushkarov
- 数海拾贝
- *Not Always Buried Deep* by Paul Pollack

Thank you!