信息安全数学基础
第六章 素性检验

中山大学 数据科学与计算机学院

1. 伪素数

根据欧拉定理, 如果n是素数, 则$\varphi(n) = n - 1$, 从而有

$$b^{n-1} \equiv 1 \bmod n$$

成立.
或者说: 如果n是素数, 则对所有与n互素的b, 都有$b^{n-1} \equiv 1 \bmod n$成立.

考虑其逆否命题: 如果存在与n互素的b使得

$$b^{n-1} \not\equiv 1 \bmod n$$

则n一定不是素数, 即n是合数.
逆否命题并不表示"如果n是合数, 则对所有与n互素的b, 都有$b^{n-1} \not\equiv 1 \bmod n$成立."

1. 伪素数

根据欧拉定理, 如果n是素数, 则$\varphi(n) = n - 1$, 从而有

$$b^{n-1} \equiv 1 \bmod n$$

成立.
或者说: 如果n是素数, 则对所有与n互素的b, 都有$b^{n-1} \equiv 1 \bmod n$成立.

考虑其逆否命题: 如果存在与n互素的b使得

$$b^{n-1} \not\equiv 1 \bmod n$$

则n一定不是素数, 即n是合数.
逆否命题并不表示"如果n是合数, 则对所有与n互素的b, 都有$b^{n-1} \not\equiv 1 \bmod n$成立."

换句话说, 当n是合数时, 只是某一个与n互素的b得$b^{n-1} \not\equiv 1 \bmod n$, 而并不是所有与$n$互素的$b$. 如果'运气'好, 恰好找到了这样的$b$, 就证实了$n$是合数.

反例: 63是合数, 8与63互素,但有$8^{62} = (2^6)^{31} \equiv 1 \bmod 63$.

对类似于63与8这样关系的数, 给一个专门的定义: 伪素数.

定义

设n是一个合数, 如果存在与n互素的整数b, 使得

$$b^{n-1} \equiv 1 \bmod n$$

成立, 则称n为*(对于基b的)*的伪素数, 有时为了区分其他性质的伪素数, 也称这种伪素数为费马伪素数*(Fermat Pseudoprime)*.

63是对于基$b = 8$的伪素数.
341, 561, 645是对于基$b = 2$的伪素数.

定理

存在无穷多个对于基2的伪素数.

要证明如果n是对于基2的拟素数，则$m = 2^n - 1$也是对于基2的伪素数.
这样，找到一个对于基2的伪素数n_0后，只需要逐次计算$n_1 = 2^{n_0} - 1, n_2 = 2^{n_1} - 1, \ldots$就得到无穷多的对于基2的伪素数.

如果n是对于基2的伪素数，那么n一定是奇合数，且$2^{n-1} \equiv 1 \bmod n$.
当n是奇合数时，存在整数d和q使得$n = dq$. 所以

$$m = 2^n - 1 = (2^d)^q - 1 = (2^d - 1)((2^d)^{q-1} + (2^d)^{q-2} + \ldots + 2^d + 1)$$

一定是合数. 如果还有$2^{m-1} \equiv 1 \bmod m$，则$m = 2^n - 1$是对于基2的一个伪素数.
因为$2^{n-1} \equiv 1 \bmod n$，所以存在整数k使得$m - 1 = 2^n - 2 = 2(2^{n-1} - 1) = kn$，
即$n \mid (m - 1)$. 因此有

$$(2^n - 1) \mid (2^{m-1} - 1),$$

即$2^{m-1} \equiv 1 \bmod m$. $\diamond$

伪素数的四个性质

1 如果n是一个奇合数，则n是对于基b的伪素数 $\iff \mathrm{ord}_n(b) | (n-1)$

"$\Longrightarrow$":

$$n\text{是对于基}b\text{的伪素数} \Longrightarrow b^{n-1} \equiv 1 \bmod n$$

伪素数的四个性质

1 如果n是一个奇合数，则n是对于基b的伪素数 $\Longleftrightarrow$ $\mathrm{ord}_n(b)|(n-1)$

"$\Longrightarrow$":

$$n\text{是对于基}b\text{的伪素数} \Longrightarrow b^{n-1} \equiv 1 \bmod n \Longrightarrow \mathrm{ord}_n(b)|(n-1)$$

"$\Longleftarrow$":

$\mathrm{ord}_n(b)|(n-1)$

伪素数的四个性质

1 如果n是一个奇合数, 则n是对于基b的伪素数 $\iff$ $\mathrm{ord}_n(b)|(n-1)$

"$\Longrightarrow$":

$$n\text{是对于基}b\text{的伪素数} \Longrightarrow b^{n-1} \equiv 1 \bmod n \Longrightarrow \mathrm{ord}_n(b)|(n-1)$$

"$\Longleftarrow$":

$$\mathrm{ord}_n(b)|(n-1) \Longrightarrow (n-1) = k\cdot \mathrm{ord}_n(b) \Longrightarrow b^{n-1} = (b^{\mathrm{ord}_n(b)})^k \equiv 1 \bmod n \quad \diamond$$

伪素数的四个性质

2 如果n是一个奇合数，如果n是对于基b_1的伪素数,且n是对于基b_2的伪素数，则n是对于基b_1b_2的伪素数

事实上，n是对于基b_1的伪素数

$$\Longrightarrow b_1^{n-1} \equiv 1 \bmod n;$$

n是对于基b_2的伪素数

$$\Longrightarrow b_2^{n-1} \equiv 1 \bmod n;$$

从而

$$\Longrightarrow (b_1b_2)^{n-1} \equiv 1 \bmod n;$$

即n是对于基b_1b_2的伪素数 $\diamond$

伪素数的四个性质

3 设n是一个奇合数，如果n是对于基b的伪素数，则n是对于基b^{-1}的伪素数

事实上，n是对于基b的伪素数

$$\Longrightarrow b^{n-1} \equiv 1 \bmod n;$$

注意到

$$\begin{aligned} b^{n-1} \cdot (b^{-1})^{n-1} &\equiv 1 \bmod n \\ \Longrightarrow (b^{-1})^{n-1} &\equiv (b^{n-1})^{-1} \bmod n \\ \Longrightarrow (b^{-1})^{n-1} &\equiv 1 \bmod n \end{aligned}$$

于是，我们有

$$b^{n-1} \equiv 1 \bmod n \Longrightarrow (b^{-1})^{n-1} \equiv 1 \bmod n.$$

即n是对于基b^{-1}的伪素数 $\diamond$

伪素数的四个性质

4 设n是一个奇合数，如果存在与n互素的整数b使得

$$b^{n-1} \not\equiv 1 \bmod n,$$

则在模n的简化剩余系

$$\{b_1, b_2, \ldots, b_s, b_{s+1}, \ldots, b_{\varphi(n)}\}$$

中至少有一半的数使得$b_i^{n-1} \not\equiv 1 \bmod n$.

假设$\{b_1, b_2, \ldots, b_s\}$使得$b_i^{n-1} \equiv 1 \bmod n$; $\{b_{s+1}, \ldots, b_{\varphi(n)}\}$使得$b_i^{n-1} \not\equiv 1 \bmod n$
则首先对b_1来说,有$(bb_1)^{n-1} \not\equiv 1 \bmod n$
因为否则的话,

$$\left.\begin{array}{r}(bb_1)^{n-1} \equiv 1 \bmod n \\ b_1^{n-1} \equiv 1 \bmod n \Longrightarrow (b_1^{-1})^{n-1} \equiv 1 \bmod n\end{array}\right\} \Longrightarrow (bb_1 \cdot b_1^{-1})^{n-1} \equiv 1 \bmod n$$

$$\Longrightarrow b^{n-1} \equiv 1 \bmod n$$

矛盾.

类似地,

$$
\begin{aligned}
(bb_2)^{n-1} &\not\equiv 1 \bmod n \\
(bb_3)^{n-1} &\not\equiv 1 \bmod n \\
&\cdots\cdots \\
(bb_s)^{n-1} &\not\equiv 1 \bmod n
\end{aligned}
$$

注意到$bb_i(i=1,2,\ldots,\varphi(n))$也是简化剩余, 这样我们找到$s$个不同的简化剩余使得

$$x^{n-1} \not\equiv 1 \bmod n.$$

又因为具有bb_i形式的简化剩余的数量一定不超过所有满足$x^{n-1} \not\equiv 1 \bmod n$的简化剩余的数量. 所以, 我们有

$$s \le \varphi(n) - s \Longrightarrow s \le \frac{\varphi(n)}{2} \Longrightarrow \varphi(n) - s \ge \frac{\varphi(n)}{2}.$$

这表明, 如果n是奇合数, 如果随机选取一个与n互素的数b, 那么$b^{n-1} \not\equiv 1 \bmod n$以较大概率($\ge \frac{1}{2}$)成立, 或者说, n是对于基b的伪素数的概率比较小($\le \frac{1}{2}$).

对于任意整数n, 如果随机选取一个与n互素的数b_i, 出现了$b_i^{n-1} \equiv 1 \bmod n$, 那么我们有理由怀疑这个数不是合数, 即这个数有较大概率是素数.

如果对n是否是素数还有担心的话, 可以继续去一个, 随机选取一个与n互素的数b_j来计算b_j^{n-1}, 看它是否与1模n同余. 如果是, 则更加确信了"n应该是一个素数" 的判断; 如果不是, 那么n一定是一个合数.

如果对n是否是素数还有担心的话, 可以继续去一个, 随机选取一个与n互素的数b_j来计算b_j^{n-1}, 看它是否与1模n同余. 如果是, 则更加确信了"n应该是一个素数" 的判断; 如果不是, 那么n一定是一个合数.

......

如果经过一定次数, 例如64次, 的检验都能确认我们的判断, 则可以说n以非常大的概率是素数.

这个检验n是否为素数的过程被称之为Fermat素性检验.

对于任意的整数n, 当然也不排除这样一种可能性, 那就是:

所有与n互素的数b, 都有$b^{n-1} \equiv 1 \bmod n$, 但n自身还是一个合数.

这种数称之为Carmichael(卡米切尔)数.

比如$n = 561$就是一个Carmichael数(因子分解$(3 \times 11 \times 17)$.

这也说明: 对所有与n互素的b都有$b^{n-1} \equiv 1 \bmod n \not\Longrightarrow n$是素数

2. Euler拟素数

根据勒让得符号和欧拉判别准则，如果$n=p$是一个素数，则对任意的整数b都有

$$b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \mod n$$

- (逆否命题成立) 给定一个整数n，如果存在与n互素的整数b使得

$$b^{\frac{n-1}{2}} \not\equiv \left(\frac{b}{n}\right) \mod n$$

  则n一定不是一个素数.

- (伪素数情况) 不排除一种情况，当n不是素数时，存在与n互素的一个整数b 使得$b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \mod n$成立. 这种$n$被称为是对基$b$的Euler伪素数.

定义

如果n是正奇合数，b与n互素且$b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \mod n$，则称$n$是对于基$b$的*Euler*伪素数.

例如，可以验证$2^{\frac{1105-1}{2}} \equiv \left(\frac{2}{1105}\right) \mod 1105$，所以1105是对于基2的Euler伪素数.

Euler伪素数的基本性质

1 n是对于基b的欧拉伪素数$\Longrightarrow$$n$是对于基$b$的伪素数.

事实上, n是对于基b的欧拉伪素数$\Longrightarrow b^{\frac{n-1}{2}} \equiv \left(\frac{b}{n}\right) \bmod n$

$$\Longrightarrow b^{n-1} \equiv \left(\frac{b}{n}\right)^2 \bmod n$$

$$\Longrightarrow b^{n-1} \equiv 1 \bmod n$$

则n是对于基b的伪素数.

这也表明, 如果n是奇合数, 且不是Carmichael(卡米切尔)数, 随机选取一个与n互素的数b, n是对于基b的伪素数的概率比较小($\leq \frac{1}{2}$), 从而n是对于基b的欧拉伪素数的概率也会比较小.

2 n是对于基b的伪素数$\not\Longrightarrow$$n$是对于基$b$的欧拉伪素数.

反例: $n = 341, b = 2.$

Solovay-Stassen素性检验

利用Euler伪素数的概念，可以判断一个给定的整数n是否是素数. 判断过程与Fermat素性检验类似，即:

1. 随机选取一个与n互素的数b_j来计算$b_j^{\frac{n-1}{2}}$和雅可比符号$(\frac{b}{n})$，并检验它们模n是否同余.
2. 如果确实同余，则可以继续下一次的检验.
3. 如果不同余，那么n一定是合数.

这个过程可以持续一定的次数，例如64次，以保证检验失误的概率足够小.

这个素性检验的方法称为Solovay-Stassen素性检验.

3. 强伪素数

设n是正奇整数，并且$n-1=2^s t$，则有下面等式成立，

$$b^{n-1}-1=(b^{2^{s-1}t}+1)(b^{2^{s-2}t}+1)\dots(b^{2^2t}+1)(b^{2t}+1)(b^t+1)(b^t-1)$$

这样，如果n是素数，b与n互素，则有$b^{n-1}-1\equiv 0 \bmod n$，从而有

$$(b^{2^{s-1}t}+1)(b^{2^{s-2}t}+1)\dots(b^{2^2t}+1)(b^{2t}+1)(b^t+1)(b^t-1)\equiv 0 \bmod n.$$

于是，在下列同余式中，

$$b^t\equiv 1 \bmod n$$
$$b^t\equiv -1 \bmod n$$
$$b^{2t}\equiv -1 \bmod n$$
$$\dots\dots$$
$$b^{2^{s-1}t}\equiv -1 \bmod n$$

至少有一个成立.

如果n是素数, b与n互素, 在下列同余式中,

$$b^t \equiv 1 \bmod n$$
$$b^t \equiv -1 \bmod n$$
$$b^{2t} \equiv -1 \bmod n$$
$$\cdots\cdots$$
$$b^{2^{s-1}t} \equiv -1 \bmod n$$

至少有一个成立.

- 逆否命题: 给定n以及与它互素的整数b, 如果上面这些同余式均不成立, 则n 一定不是素数;
- 伪素数情况: 如果n是合数, 有可能存在整数b使得在上面这些同余式中至少有一个成立, 这种n被称为是关于b的强伪素数.

定义

设n是一个奇合数, $n-1=2^s t$, 其中t为奇数. 对于与n互素的整数b, 如果存在一个整数i, 满足$0 \le i < s$, 使得

$$b^{2^i t} \equiv -1 \bmod n,$$

则称n是对于基b的强伪素数(*Strong Pseudoprime*).

相反地, 如果

$$\begin{aligned} b^t &\equiv 1 \bmod n \\ b^t &\equiv -1 \bmod n \\ b^{2t} &\equiv -1 \bmod n \\ &\cdots\cdots \\ b^{2^{s-1}t} &\equiv -1 \bmod n \end{aligned}$$

都不成立, 这就意味着n一定不是素数. 因为, 如果n是素数, 则上面这些同余式至少有一个成立. 这样我们也就得到了类似前面判断素数的方法.

Miller-Rabin素性检验

1. 将$n-1$写成$n-1=2^s t$, 其中t为奇整数.
2. 随机取整数b,
3. 计算$r_0=b^t \bmod n$, 如果$r_0=\pm 1$, 则判断n是素数;
4. 如果$r_0\neq \pm 1$, 计算$r_1=r_0^2\equiv b_0^{2t} \bmod n$, 如果$r_1=-1$, 则判断$n$是素数;
5. 如果$r_1\neq -1$, 计算$r_2=r_1^2\equiv b_0^{2^2 t} \bmod n$, 如果$r_2=-1$, 则判断$n$是素数;
6. 如果$r_2\neq -1$, 计算$r_3=r_2^2\equiv b_0^{2^3 t} \bmod n$, 如果$r_3=-1$, 则判断$n$是素数;
7. 重复上述计算过程, 如果一直有$r_i\neq -1 (1\leq i<s-1)$,
   计算到$r_{s-1}=r_{s-2}^2\equiv b_0^{2^{s-1} t} \bmod n$, 如果$r_{s-1}=-1$, 则判断$n$是素数;
8. 否则, 对于$1\leq i<s$, 如果都有$r_i\neq -1$, 则判断n是合数.

这个检测方法称为Miller-Rabin素性检验.

强伪素数的四个性质

1 存在无穷多个对于基2的强伪素数.

将证明"n是对于基2的伪素数，则$m=2^n-1$是对于基2的强伪素数."
由于在无穷多个对于基2的伪素数，因此，也就存在无穷多个对于基2的强伪素数.

回忆伪素数的性质，如果n是对于基2的伪素数，则$m=2^n-1$必定是奇合数.另外，如果n是对于基2的伪素数，则$2^{n-1}\equiv 1 \bmod n$，即存在奇数k使得$2^{n-1}-1=nk$.

根据强伪素数的定义，需要考虑$m-1$的分解形式，

$$m-1=2^n-1-1=2^n-2=2(2^{n-1}-1)=2(nk).$$

注意到

$$2^t=2^{nk}=(2^n)^k=(2^n-1+1)^k=(m+1)^k,$$

从而有$2^t\equiv 1 \bmod m$，即$m=2^n-1$是对于基2的强伪素数.

强伪素数的四个性质

2 n是对于基b的强伪素数，则n是对于基b的欧拉伪素数，从而n是对于基b的伪素数.

这个性质表明，如果n是奇合数，且不是Carmichael数，随机选取一个与n互素的数b，n是对于基b的伪素数的概率比较小($\leq \frac{1}{2}$)，从而n是对于基b的欧拉伪素数的概率也会比较小，从而n是对于基b的强伪素数的概率也会比较小.

3 n是奇合数. 随机选取与n互素的数b，则n是对于基b的强伪素数的概率至多为$1/4$.
这个性质表明，Miller-Rabin素性检验比Solovay-Stassen素性检验更好.

4 不存在一个Carmichael数n，对于任意一个与它互素的数b，它是对于基b的欧拉伪素数(强伪素数).
这个性质表明，Solovay-Stassen素性检验，以及Miller-Rabin素性检验，有可能检验出某一个Carmichael数是合数.