初等数论及其应用

第二章 同余

2.9 伪素数和素性测试

素性测试的讨论

- 在前面的章节中，已经得到了几种素性测试的方法，如**试除法**、**威尔逊定理**等：
  - 在实际应用中，考虑到计算量的问题，试除法仅仅适用于12到14位数字的数.
  - 威尔逊定理表明，一个整数n是素数的充要条件是$(n-1)! \equiv -1 \ (mod\ n)$，因此理论上讲，只要计算出$(n-1)!$，就容易判断$n$是否是素数了. *但当$n$较大时无有效方法计算阶乘*，故威尔逊定理也不适用于对较大的数作素性测试.

- 此外，根据**费马小定理**，若p是素数，则对任意整数a都有$a^p \equiv a \ (mod\ p)$. 因此，若存在某个整数a使得$a^n \not\equiv a \ (mod\ n)$，则$n$肯定不是素数. 这里，**计算$a^n$关于模$n$的余数**，可利用模算术的技巧**在多项式时间内完成**(以后将介绍).

素性测试的讨论

- 关于费马小定理，下面问题值得考虑：如果$a^n \equiv a \,(mod\, n)$对所有整数a都成立，那么n是否是素数？如果这样的n不是素数，那么这样的数多吗？它们有何特征？

- 如果想找到满足$a^n \not\equiv a \,(mod\, n)$的$a$，从而判定$n$是合数，那么可以从最简单的情形$a = 2$开始测试.

  如果某个a使得$a^n \equiv a \,(mod\, n)$，那么称n**通过基a测试**；否则称n**未通过基a测试**，此时n必为合数.

  计算发现，当$1 < n < 341$时，所有通过基2测试的都是素数.

  不过当$n = 341$时，

  $2^{341} \equiv (2^{10})^{34} \times 2 \equiv 1024^{34} \times 2 \equiv 1^{34} \times 2 \equiv 2 \,(mod\, 341)$

  即341也通过了基2测试，但是$341 = 11 \times 31$是合数.

  因此，通过基2测试的数并非都是素数.

基2的伪素数

- **定义2.8.1** 如果一个**合数n**满足$2^n \equiv 2 \ (mod\ n)$, 则称n是一个**基2的伪素数**.

- 由上述定义, 341是一个基2的伪素数. 事实上, 基2的伪素数有无穷多个, 即是下面的定理:

- **定理2.8.1** 存在无穷多个基2的伪素数.

- **证明**: 只需证明, **如果n是一个基2的伪素数, 则$m = 2^n - 1$也是一个基2的伪素数**. 因为341是一个基2的伪素数, 所以该断言表明基2的伪素数有无穷多个. 下面将证明这个断言.

基2的伪素数

- **证明:** 由n是基2的伪素数知, n是合数, 且$2^n \equiv 2 \ (mod\ n)$.

  不妨设为$n = st \ (s, t > 1)$, 则由

  $\boldsymbol{m = 2^n - 1 = (2^s - 1)(2^{s(t-1)} + \cdots + 2^s + 1)}$

  知m也是合数.

  因为$2^n \equiv 2 \ (mod\ n)$, $m - 1 = 2^n - 2$, 所以$n | m - 1$.

  令$m - 1 = kn$, 则有

  $$2^{m-1} - 1 = 2^{kn} - 1 = (2^n - 1)\left(2^{n(k-1)} + \cdots + 2^n + 1\right)$$

  $$= m(2^{n(k-1)} + \cdots + 2^n + 1)$$

  故$m | 2^{m-1} - 1$, 即$m | 2^m - 2$, 故m也是一个基2的伪素数.

- 上述证明表明, **存在无穷多个基2的奇伪素数.**
- 1950年, 莱梅发现了第一个基2的偶伪素数. 事实上, **基2的偶伪素数也有无穷多个.**

基a的伪素数

- 作为基2的伪素数的推广，我们下面将考虑更一般的情形，即基a的伪素数.

- **定义2.8.2** 设整数$a > 1$，如果一个合数n满足$a^n \equiv a(mod\ n)$，则称n是一个**基a的伪素数**.

- 显然，如果$(a,n) = 1$，那么上述定义中的条件等价于 $a^{n-1} \equiv 1\ (mod\ n)$.

- **定理2.8.2** 对任意整数$a > 1$，均存在无穷多个基a的伪素数.
- 证明略（留作思考题），感兴趣的同学可以参见教材.

费马素性测试

- 上述定理说明，整个整数集中存在无穷多个以a为基的伪素数. **但是**，在任意给定的范围内，**伪素数的数目比素数要少得多**.

  例如，小于10^9的素数有50847534个，但小于10^9的基2的伪素数只有5597个.

- 因此，对随机选取的小于10^9的正整数n，如果$2^n \equiv 2(mod\ n)$，那么n是素数的可能性将非常大. 事实上，n为素数的概率 $50847534/(50847534+5597) > 99.9\%$. 如果使用这种方法来判断$n$是否是素数，那么出错的概率将很小.

  另外，还可以选择多个基对n测试，进一步降低出错的概率. 这种基于费马小定理的素性测试方法通常称为**费马素性测试**，它显然是一种概率算法.

卡米歇尔数的定义

- 人们自然希望对每个合数n, 都能通过不断更换基a, 从而使得n不能通过某个基a测试. 然而, 事实表明, 存在合数n, 它能通过任意基a测试, 卡米歇尔最早研究了这样的数.

- **定义2.8.3** 设n是一个合数, 如果对每个正整数a, 都有$a^n \equiv a(mod\ n)$, 则称n是一个**卡米歇尔数**或**绝对伪素数**.

- 1910年, 卡米歇尔给出了15个**卡米歇尔数**, 并猜想这样的数有无穷多个. 这个猜想在1994年得到证明. 最初的10个卡米歇尔数分别是: 561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341.

卡米歇尔数的定义

- **例2.8.1** 证明561是卡米歇尔数.

- **证明:** 因为$561 = 3 \times 11 \times 17$,

  所以对每个正整数a验证$a^{561} \equiv a \ (mod\ 561)$, 只需分别验证:

  $a^{561} \equiv a \ (mod\ 3), a^{561} \equiv a \ (mod\ 11), a^{561} \equiv a \ (mod\ 17)$

  1) 若$3|a$, 则显然$a^{561} \equiv a \ (mod\ 3)$, 而若$3 \nmid a$, 则由费马小定理知

  $a^2 \equiv 1(mod\ 3)$, 故$a^{561} \equiv (a^2)^{280}a \equiv a \ (mod\ 3)$.

  2) 若$11|a$, 则显然$a^{561} \equiv a \ (mod\ 11)$, 而若$11 \nmid a$, 则$a^{10} \equiv 1 \ (mod\ 11)$,

  于是$a^{561} \equiv (a^{10})^{56}a \equiv a \ (mod\ 11)$.

  3) 若$17|a$, 则显然$a^{561} \equiv a \ (mod\ 17)$, 而若$17 \nmid a$, 则$a^{16} \equiv 1 \ (mod\ 17)$,

  于是$a^{561} \equiv (a^{16})^{35}a \equiv a \ (mod\ 17)$.

  由于上述三式同时成立, 故$a^{561} \equiv a \ (mod\ 561)$. 得证!

卡米歇尔数的性质

- **命题2.8.1** 设n是卡米歇尔数, 则: (1) n是奇数; (2) n是不同素数之积.

- **证明:** (1) 因为n是卡米歇尔数,

  所以当$a = n - 1$时, 有$(n-1)^n \equiv n - 1 \equiv -1\ (mod\ n)$,

  而$(n-1)^n \equiv (-1)^n\ (mod\ n)$, ---(展开$(n-1)^n$即可)

  故$(-1)^n \equiv -1\ (mod\ n)$. 于是$n = 2$或$n$为奇数.

  因为卡米歇尔数是合数, 所以n为奇数.

  (2) 假设存在素数p使得$p^2|n$.

  因为n是卡米歇尔数, 所以当$a = p$时, $p^n \equiv p\ (mod\ n)$, 即$n|p^n - p$.

  因此$p^2|p^n - p$, 从而$p^2|p$ --(这里用到$p^2|p^n$),

  产生矛盾. 故n是不同素数之积.

卡米歇尔数的Korselt判定方法

- **定理2.8.3** 设n是合数, 则n是卡米歇尔数当且仅当整除n的每个素数p满足: (1) $p^2 \nmid n$; (2) $p-1|n-1$.

- **证明**: 先证**充分性**.

  由条件1, 可设$n = p_1 p_2 \dots p_k$, 其中$p_1, p_2, \dots, p_k$是不同的素数.

  由条件2, 得$p_i - 1|n-1$, 设$n - 1 = (p_i - 1)s_i, i = 1,2, \dots, k$.

  对任意正整数a, 如果$p_i|a$, 则显然$a^n \equiv a \ (mod\ p_i)$;

  如果$p_i \nmid a$, 则由费马小定理知$a^{p_i-1} \equiv 1(mod\ p_i)$,

  从而$a^n = a^{(p_i-1)s_i+1} = (a^{(p_i-1)})^{s_i} \cdot a \equiv a(mod\ p_i), i = 1,2, \dots, k$.

  因为$[p_1, p_2, \dots, p_k] = p_1 p_2 \dots p_k = n$, 所以$a^n \equiv a \ (mod\ n)$.

  故n是卡米歇尔数.

  再证**必要性**.

  由命题2.8.1的证明即得$p^2 \nmid n$,

  而$p-1|n-1$的证明需要用到群论或原根知识, 在此略过.

卡米歇尔数的定义

- **例2.8.1** 证明561是卡米歇尔数.

- **证明:** 因为$561 = 3 \times 11 \times 17$,

  1) $3^2 \nmid 561$, 且$(3-1)|(561-1)$.

  2) $11^2 \nmid 561$, 且$(11-1)|(561-1)$.

  3) $17^2 \nmid 561$, 且$(17-1)|(561-1)$.

  由于上述三式同时成立,故561是卡米歇尔数. 得证!

卡米歇尔数的性质

- **推论2.8.1** 设n是卡米歇尔数, 则n至少有3个不同的素因数.

- **证明:** 因为n是合数, 且由命题2.8.1知, n没有平方因数,
  所以可用反证法.
  假设$n = p_1p_2$, 其中p_1, p_2是不同的素数.
  由定理2.8.3, $p_1 - 1 | n - 1$.
  因为$n - 1 = (p_1 - 1)p_2 + p_2 - 1$, 所以$p_1 - 1 | p_2 - 1$.
  同理可得$p_2 - 1 | p_1 - 1$, 因此$p_1 = p_2$与假设矛盾.
  故n至少有3个不同的素因数.

- 我们已经知道, 最小的卡米歇尔数561是3个不同的素数之积. 上面推论说明, 不存在比561有更少素因数的卡米歇尔数.

米勒-拉宾素性测试

- 卡米歇尔数存在，大大降低了费马素性测试的准确性. 下面介绍更好的素性测试方法，它基于素数如下性质：

- **定理2.8.4** 设p是奇素数，令$p-1=2^{\alpha}t$，其中t是奇数. 如果整数a满足$p \nmid a$，那么下述**两个条件之一**成立：

  (1) $a^{t} \equiv 1 \ (mod\ p)$

  (2) 存在$0 \leq i \leq \alpha - 1$，使得$a^{2^{i}t} \equiv -1 \ (mod\ p)$

- 如果存在a，使得正奇数n没有上面定理中素数p所具有的性质，那么n必是合数. 如果对a的许多不同取值，n都具有上面性质，那么n很可能是素数. 这种方法就是**米勒-拉宾素性测试**.

米勒-拉宾素性测试

- **证明:** 因为$p \nmid a$, 所以$a^{p-1} \equiv 1(mod\ p)$, 即$a^{2^{\alpha}t} \equiv 1(mod\ p)$,

  也即$(a^{2^{\alpha-1}t})^2 \equiv 1(mod\ p)$,

  由于同余方程$x^2 \equiv 1(mod\ p)$恰好有两个解$x \equiv \pm 1(mod\ p)$,

  因此有 $a^{2^{\alpha-1}t} \equiv 1(mod\ p)$或$a^{2^{\alpha-1}t} \equiv -1(mod\ p)$

  显然这两个同余式不可能同时成立,

  否则有$-1 \equiv 1(mod\ p)$, 与p为奇素数矛盾!

  如果第二个同余式成立, 那么条件2成立;

  否则, 有$a^{2^{\alpha-1}t} \equiv 1(mod\ p)$, 也即$(a^{2^{\alpha-2}t})^2 \equiv 1(mod\ p)$,

  重复上述对$(a^{2^{\alpha-1}t})^2 \equiv 1(mod\ p)$类似的讨论,

  直到发现某个$0 \le i \le \alpha - 1$, 使得$a^{2^i t} \equiv -1(mod\ p)$.

  如果对 所有$0 \le i \le \alpha - 1$都不满足$a^{2^i t} \equiv -1(mod\ p)$,

  那么进行到$i = 0$时, 必有$a^t \equiv 1(mod\ p)$. 定理得证!

米勒-拉宾素性测试

- **定理2.8.5 (米勒-拉宾素性测试)** 设n是大于1的正奇数, 令$n-1=2^{\alpha}t$, 其中t是奇数. 如果对满足$n \nmid a$ 的某个整数a, 下述两个条件都成立, 那么n必是合数:

  (1) $a^{t} \not\equiv 1 \ (mod\ n)$

  (2) 对所有$0 \leq i \leq \alpha-1$, 都有$a^{2^{i}t} \not\equiv -1 \ (mod\ n)$

- 上述米勒-拉宾素性测试的结论, 可直接由**定理2.8.4**推出.
- 利用米勒-拉宾素性测试, 我们可以验证最小卡米歇尔数561是合数, 这是因为$561-1=2^{4}\times 35, \alpha=4, t=35$, 且对$a=2$, (显然$561 \nmid 2$)有:

  $2^{35} \equiv 263 \not\equiv 1 \ (mod\ 561)$; $2^{35} \not\equiv -1 \ (mod\ 561)$;

  $2^{70} \equiv 166 \not\equiv -1 \ (mod\ 561)$; $2^{140} \equiv 67 \not\equiv -1 \ (mod\ 561)$;

  $2^{280} \equiv 1 \not\equiv -1 \ (mod\ 561)$.

米勒-拉宾素性测试

- 对正奇数$n = 2^{\alpha}t + 1$, 如果有满足$n \nmid a$的某个正整数a, 使得$a^{t} \equiv 1\ (mod\ n)$或者存在$0 \leq i \leq \alpha - 1$满足$a^{2^{i}t} \equiv -1\ (mod\ n)$, 那么称$n$**通过基$a$米勒-拉宾素性测试**, 否则称$n$**未通过基$a$米勒-拉宾素性测试**.

- 如果n通过基a米勒-拉宾素性测试, 那么或者$a^{t} \equiv 1\ (mod\ n)$或者存在$0 \leq i \leq \alpha - 1$满足$a^{2^{i}t} \equiv -1\ (mod\ n)$.

  无论哪种情况, 因为对任意$0 \leq \beta \leq \alpha - 1$, 都有$a^{n-1} = (a^{2^{\beta}t})^{2^{\alpha-\beta}}$,

  所以$a^{n-1} \equiv 1\ (mod\ n)$, 即$a^{n} \equiv a\ (mod\ n)$,

  故n通过基a测试, 即n可能为伪素数.

  因此, 也称通过基a米勒-拉宾素性测试的合数为**基a的强伪素数**.

- 例如: 合数$n = 2047 = 23 \times 89$是基2的伪素数, 也是基2的强伪素数

米勒-拉宾素性测试

- 下面的定理表明，不存在类似卡米歇尔数的强伪素数，即不存在奇合数n能够通过所有的基a米勒-拉宾测试，其中a为任意大于1的正整数.

- **定理2.8.6** 如果n是奇合数，那么n至多通过$(n-1)/4$个基a取值在$\{1,2,\ldots,n-1\}$中的米勒-拉宾测试.

- **证明：** 过程比较复杂，有兴趣的同学参见[K. H. Rosen. Elementary Number Theory and Its Applications (5th Edition). Reading, MA: Addison-Wesley, 2005]

- 由上述定理知，如果一个奇数n通过多于$(n-1)/4$个基a取值在$\{1,2,\ldots,n-1\}$中的米勒-拉宾测试，那么n必是素数. 这种确定性素性测试方法在实际应用中计算量太大. 但是，通过选取少量的基，米勒-拉宾测试能给出一种快速判断一个数可能是素数的概率算法.

米勒-拉宾素性测试

- 由定理2.8.6知，n通过基a米勒-拉宾测试的概率小于1/4. 如果在$\{1,2,\ldots,n-1\}$中随机选取k个不同的基分别进行米勒-拉宾测试，那么n通过所有k个基的米勒-拉宾测试的概率将小于$1/4^k$. 例如，对奇合数n，如果随机选取$k=100$个1到n间的整数作为基对n进行米勒-拉宾测试，那么n能通过这100个测试的概率将小于10^{-60}.

- 1976年，基于解析数论中著名的(尚未证明的)广义黎曼假设，米勒证明了下面事实：如果广义黎曼假设成立，那么对每个合数n，都存在基$a, 1<a<2(\log_2 n)^2$，使得n不能通过基a米勒-拉宾测试.

总结

- **定义2.8.2** 设整数$a > 1$, 如果一个合数n满足$a^n \equiv a(mod\ n)$, 则称n是一个**基a的伪素数**.
- **定理2.8.2** 对任意整数$a > 1$, 均存在无穷多个基a的伪素数.
- **定义2.8.3** 设n是一个合数, 如果对每个正整数a, 都有$a^n \equiv a(mod\ n)$, 则称n是一个**卡米歇尔数**或**绝对伪素数**.
- **定理2.8.3** 设n是合数, 则n是卡米歇尔数当且仅当整除n的每个素数p满足: (1) $p^2 \nmid n$; (2) $p - 1 | n - 1$.
- **推论2.8.1** 设n是卡米歇尔数, 则n至少有3个不同的素因数.

总结

- **定理2.8.5 (米勒-拉宾素性测试)** 设n是大于1的正奇数, 令$n-1=2^{\alpha}t$, 其中t是奇数. 如果对满足$n \nmid a$ 的某个整数a, 下述两个条件之一都成立, 那么n必是合数:

  (1) $a^{t} \not\equiv 1 \ (mod\ n)$

  (2) 对所有$0 \leq i \leq \alpha-1$, 都有$a^{2^{i}t} \not\equiv -1 \ (mod\ n)$

- 通过基a米勒-拉宾素性测试的合数为**基a的强伪素数**.

- **定理2.8.6** 如果n是奇合数, 那么n至多通过$(n-1)/4$个基a取值在$\{1,2,\ldots,n-1\}$中的米勒-拉宾测试.