暗号の数理要綱 #9

4.4 素数判定

RSA 暗号においては，100〜200 桁程度の (ランダムな) 素数を 2 つ用意する必要がある。RSA 暗号の安全性は，大きな数を素因数分解することは非常に難しい，という事実に基づいているが，大きな素数を用意するためには，その数が素数であることを確かめなければならず，そのためには，その数が素因数分解できない，ということを示さなければならない，という矛盾が生じてしまう。

この，一見，致命的とも思えるような矛盾を巧妙に回避する方法が **素数判定** である。素数判定とは，

- 「ある数が素数ではない」ということを示す方法

のことである。

「ある数が素数である」ことを示すような，どの数にでも使える効率的な方法は知られていない。「素数ではない」ことを示すような判定法を何回適用しても「素数ではない」ということを示すことができなかったとしたら，その数は，かなり高い確率で素数らしい，ということになる。

RSA 暗号においては，そのような数を素数であると見なして，素数であるとして使う。万が一，それが素数ではなかったとしたら，復号化の過程で問題が発生し，元の平文に復号化できなくなるかもしれないが，その確率が非常に小さければ問題なかろう，という発想である。

4.4.1 フェルマー・テスト

n という自然数を一つ固定し，これが素数であるかどうかを判定することを考える。

n の倍数ではない自然数 a があって

$$a^{n-1} \not\equiv 1 \mod n \qquad (*)$$

となったとしよう。もし n が素数なら，フェルマーの小定理より，このようなことは起こり得ない。従って，もしこのようなことが成立するならば,「 n は素数ではない 」 ということの証明になる。

特に，$1 < a < n$ という自然数 a は n の倍数となることはないから，このようなある a で $(*)$ のようなことになれば，因数分解をすることなく，n は合成数であることが証明できることになる。

任意の自然数 b は，ある自然数 k と，$0 \leqq r \leqq n-1$ となる自然数 r があって $b = kn + r$ と書ける。2 項展開すると，ある自然数 q があって，

$$b^{n-1} = (kn+r)^{n-1} = r^{n-1} + \binom{n-1}{1} r^{n-2} kn + \cdots + (kn)^{n-1} = r^{n-1} + nq$$

となることがわかるから，

$$b^{n-1} \equiv r^{n-1} \mod n$$

である。従って，このフェルマーの小定理を使って，n が素数ではないことを示そうとした場合，a として，$2 \leqq a \leqq n-1$ となるものだけを考えれば十分であることがわかる。

また，$2 \leqq a \leqq n-1$ となる a に対して $GCD(a,n) \neq 1$ であるなら，n は素数ではあり得ない。$GCD(a,n)$ はユークリッド互除法で短時間で計算できるから，フェルマーの小定理を適用する前に，まず，この判定法行ってみることは時間の節約になる。

定義 4.1 n を合成数とする。$1 < a < n$ かつ $GCD(a,n) = 1$ となるある a に対して

$$a^{n-1} \equiv 1 \mod n$$

となる時，n は a を底とする **擬素数** であるという。

これは，n は合成数であるにも関わらず，a を使う限りにおいて，合成数であることを示すことはできない，という意味である。
このように，$1 < a < n$ となる a に対して

(1) $GCD(a,n)$ を計算し，

(2) $a^{n-1} \equiv 1 \mod n$ となるのかどうかを試す

ことを **フェルマー・テスト** という。

このフェルマー・テストは，実は意外なほど強力である。

例えば，1〜1,000 までの中に素数は 168 個 ある。これは，残りの 832 個 は合成数であるということを意味する。この中で，2 を底とする擬素数は，

$$341, \quad 561, \quad 645$$

の 3 個しかない。すなわち，この 3 個以外の 829 個 の合成数は全て，$a = 2$ **だけ** を使うことにより，フェルマー・テストで合成数であることを証明できる，ということになる。

なお，この 3 個の合成数は，3 を底とする **擬素数ではない** ので，2 と 3 だけを使えば，1,000 までの全ての合成数を，フェルマー・テストで判定できる。

また，10,000 までの全ての合成数は，2，3，5，7 だけを使って，フェルマー・テストにより，合成数であることを判定できる。100,000 までの合成数で，この 4 個で判定できないものは，

$$29341, \quad 46657, \quad 75361$$

の 3 個だけである。

実際にこのフェルマー・テストの対象となる n は非常に大きな数なので，小さな a から順に試して行く，という方法を取ることはできない。$1 < a < n$ を満たす a をランダムにとり，何回か試して，合成数であることを証明できなければ，素数である確率が高い，と見なすわけである。

定義 4.2 n を奇数の自然数で，さらに合成数であるとする。$2 \leqq a \leqq n-1$ となる n **と互いに素な** a に対して $a^{n-1} \equiv 1 \pmod{n}$ となる時，n を **カーマイケル数** と言う。

n がカーマイケル数なら，ランダムに選ばれた $1 < a < n$ が $GCD(a,n) = 1$ なら（そうなる確率は高い），その a では，フェルマーの小定理で n が合成数であることを示すことはできない，ということである。

$561 = 3 \cdot 11 \cdot 17$ はカーマイケル数であることが知られている。また，カーマイケル数は無限に存在することも知られている。カーマイケル数がどのようなものであるのかは，以下の定理によって，良くわかっている。

定理 4.3 奇数の合成数 $n \geqq 3$ がカーマイケル数であるための必要十分条件は，次の 2 つの条件が成り立つことである。

(1) n の素因数は平方以上を含まない。すなわち，p が n の素因数なら，p^2 は n の約数ではない。

(2) p が n の素因数なら，$p-1$ は $n-1$ の約数である。

4.4.2 Miller-Rabin 法

素数判定を行うためのフェルマー・テストは，カーマイケル数が存在するという理由から，もちろん完全な判定法ではないが，それ以上に，実際に試してみると，許容できる計算時間の範囲内では，無視できない割合で判定ミスが起こることがわかる。

実際に使用するためには，より効率的で，かつ精度の高い判定法が必要となる。ここでは，実際に素数判定に使われている Miller-Rabin 法について述べる。

素数判定法とは，

「素数が必ず持っていなければならない性質A」

というものを一つ選ぶことである。そして，ある数がこの「性質A」を持っていない，ということを示すことができれば，その数は素数ではない，ということを決定できることになる。

例えばフェルマー・テストでは,「フェルマーの小定理が成り立つ」というのが，ここでの「性質A」ということになり，フェルマーの小定理が成立しない数は合成数である，と判定できる。

素数が満たすべき性質というのは無数にあり，どれが判定法として適切か，というのは，非常に難しい問題である。

1976 年，G.L.Miller が提案したのは次のような性質である。

$n \geqq 3$ を素数であると仮定する。

$n-1$ は偶数なので，2 で割り切れる。2 の素因数を全て抜き出すことにより，q を奇数として $n-1 = 2^k q$ と書ける。ここで $k \geq 1$ である。

n は素数と仮定したので，フェルマーの小定理より $0 < b < n$ となる任意の b に対して

$$b^{2^k q} = b^{n-1} \equiv 1 \mod n$$

である。そこで，

$$b^q, \quad b^{2q}, \quad b^{2^2 q}, \quad \cdots \quad b^{2^{k-1} q}, \quad b^{2^k q}$$

という $k+1$ 個の数を $\mod n$ で考える。右端のものは 1 となるので，この中に 1 があることは保障されている。そこで $j \geq 0$ を，$b^{2^j q} \mod n = 1$ となる最小の非負整数とする。

$\left(b^{2^i q}\right)^2 = b^{2^i q \cdot 2} = b^{2^{i+1} q}$ なので，どこかで 1 が現れたら，その右側は全て 1 であることに注意。特に $j = 0$ ならば，全てが 1 である。

さて，$j \geq 1$ とする。$b^{2^j q} - 1 \equiv 0 \mod n$ であるが，

$$b^{2^j q} - 1 = \left(b^{2^{j-1} q} - 1\right)\left(b^{2^{j-1} q} + 1\right) \equiv 0 \mod n$$

と因数分解してみると，j は $b^{2^j q} - 1 \mod n = 0$ となる最小のものだったので，$b^{2^{j-1} q} - 1 \neq 0 \mod n$ である。従って $b^{2^{j-1} q} + 1 = 0 \mod n$ でなければならない。すなわち，

$$b^{2^{j-1} q} \equiv -1 \mod n$$

である。このとき n が素数であるという性質を使っていることに注意すること。n が素数でなければ，$b^{2^j q} - 1 \equiv 0$ と $b^{2^{j-1} q} - 1 \not\equiv 0 \mod n$ から $b^{2^{j-1} q} + 1 \equiv 0$ はでてこない。

これが意味することをまとめると次のようになる。

命題 4.4 $n \geq 3$ を素数とし $1 < b < n$ とする。また，q を奇数として $n - 1 = 2^k q$ であるとする $(k \geq 1)$。この時，$b^q = 1 \mod n$ であるかまたは，$b^{2^r q} \equiv -1 \mod n$ となる $0 \leq r \leq k - 1$ が存在する。

これが，素数が満たさなければならない一つの性質であり，これを満足しない n は「素数ではない」と判定できることになる。具体的な判定手続きは次のようになる。

（なおここで，$-1 \equiv n - 1 \mod n$ であることに注意。すなわち，実際の計算では -1 ではなく $n - 1$ が現れることが普通である。）

(1) 奇数 $n \geq 3$ を一つとる。これが素数であるかどうか判定したい。

(2) q を奇数，$k \geq 1$ として $n - 1 = 2^k q$ とする。

(3) $1 < b < n$ となる b を一つ選ぶ。

(4) もし $(b, n) \neq 1$ ならば n は素数ではないと判定できる。以下，$(b, n) = 1$ であったとする。

(5) $b^q \mod n = \pm 1$ ならば，n は b を使う限りにおいては「素数ではない」と判定できない。すなわち，素数である可能性がある。別の b を選ぶために (3) に戻る。

(6) $b^q \mod n \neq \pm 1$ の場合，$j \geq 1$ に対して $b^{2^j q} \mod n$ を計算して行くのだが，$\left(b^{2^i q}\right)^2 = b^{2^{i+1} q}$ なので，$b^q \mod n = m_0$ として，$m_{j+1} = (m_j)^2 \mod n$ を計算して行けば良い。

(7) ある $1 \leq j \leq k - 1$ で $m_j = -1$ となったら，n は b を使う限りにおいては「素数ではない」と判定できない。すなわち，素数である可能性がある。別の b を選ぶために (3) に戻る。

(8) 全ての $1 \leq j \leq k - 1$ で $m_j \neq -1$ ならば，命題 4.4 より，そのようなことは素数では起こり得ないことなので，n は合成数であると判定できる。

カーマイケル数はフェルマー・テストでは合成数であることを判定できなかった。最小のカーマイケル数 $561 = 3 \cdot 11 \cdot 17$ に関して，この Miller 法を試してみよう。

$561-1=560=2^4\cdot 35$ であるから，$1<b<560$ であり $(b,561)=1$ となる b に関して，

$$b^{35} \mod 561, \quad b^{2\cdot 35} \mod 561, \quad b^{2^2\cdot 35} \mod 561, \quad b^{2^3\cdot 35} \mod 561$$

を計算してみる。

$b=2$ を採用してみよう。$35=32+3=2^5+3$ であるから，$2^{35}=2^{2^5+3}=2^{2^5}\cdot 2^3$ である。高速指数計算法より，

$$\begin{aligned} &2^2 \mod 561 = 4, && 4^2=2^{2^2} \mod 561 = 16,\\ &16^2=2^{2^3} \mod 561 = 256, && 256^2=2^{2^4} \mod 561 = 460,\\ &460^2=2^{2^5} \mod 561 = 103, && 103\cdot 2^3 = \mod 561 = 263 \end{aligned}$$

であるから，$2^{35} \mod 561 = 263$ である。

$$\begin{aligned} 263^2 \mod 561 = 2^{2\cdot 35} \mod 561 &= 166\\ 166^2 \mod 561 = 2^{2^2\cdot 35} \mod 561 &= 67\\ 67^2 \mod 561 = 2^{2^3\cdot 35} \mod 561 &= 1 \end{aligned}$$

となり -1 は出てこない。従って，底として 2 を使うだけで，561 は合成数であると判定できる。

定義 4.5 b を固定した時に，このテストをパスしてしまう合成数 n を，b を底とする **強擬素数** という。

強擬素数はいくらでも存在する。例えば 25 は，$25-1=24=2^3\cdot 3$ であり，$(7,25)=1$ なので，7 を底として Miller テストを行ってみると，

$$7^3 \mod 25 = 18, \quad 18^2=7^{2\cdot 3} \mod 25 = 24 = -1 \mod 25$$

となり強擬素数となる。

しかし，今度は 2 を底として Miller テストを行ってみると，

$$2^3 \mod 25 = 8, \quad 8^2=2^{2\cdot 3} \mod 25 = 14, \quad 14^2=2^{2^2\cdot 3} \mod 25 = 21$$

となり -1 は出てこないので，合成数であると判定できる。

この Miller テストがフェルマー・テストと根本的に異なる点は,「カーマイケル数的」な数が存在しない，ということである。これは，1980 年に M.O.Rabin によって示された。

定理 4.6 $n\geq 3$ を奇数とする。$n/4$ 個よりも多い $1<b<n$ に対して n が Miller テストをパスするならば，n は素数である。

すなわち，全ての底について，テストをすり抜けてしまうような合成数は存在しない，ということである。これは実に素晴らしい結果ではあるが，現実的には，10^{100} 程度の大きな n に関して $n/4$ 個の底について Miller テストを行うことは不可能なので，完全な判定はできない。(例えば

$n = 10^{100}$ の場合，$n/4$ は 10^{99} よりも大きいことに注意。$n/4 \approx 10^{25}$ などと錯覚しないように。$10^{25} = n/10^{75}$ である。もちろん 10^{25} でさえ，手に負えないほど大きな数である。)

しかしながら，定理 4.6 の対偶をとると，

「n が合成数ならば，$n/4$ 個よりも多い b を底として強擬素数となることはない。」

ということであるから，n が合成数であるとすると，ランダムに選んだ $1 < b < n$ と $(b, n) = 1$ を満たす b に関して，n が強擬素数となる確率は $1/4$ 以下である。さらに，r 個のランダムに選んだ b 全てに関して強擬素数となる確率は $(1/4)^r$ 以下である。

例えば $r = 20$ なら，$(1/4)^{20} \approx 1/10^{12}$ であるから，合成数が，20 個のランダムに選んだ底に対して強擬素数となる確率は「1 兆分の一」ということになる。ということは，20 個のランダムに選んだ底に対して Miller テストをパスするような数については,「素数である確率が非常に高い」と言えることになる。

ただしこれはあくまで確率であって，いくら小さくても 0 にはならない。完全に素数であることを保障するためには，$n/4$ 個よりも多い b に関して Miller テストをパスすることを示さなければならないことを注意しなければならない。

4.4.3 Miller-Rabin 法の計算量

Miller-Rabin 法は，上記の判定確率が評価できるという優れた点を持っていると同時に，計算量が少ない，大きな利点も持ち合わせている。

$n - 1 = 2^k q$ とした時，k という数の大きさが問題となるが，この k は最大で $n \approx 2^k$ となるくらいの数であり，およそ $\log_2 n$ 程度である。

例えば $n = 10^{200}$ ならば $\log_2 10^{200} \approx 664$ 程度の数である。1 回の計算は，2 乗して $\mod n$ をとるというものであり，それを最大で 600 回程度行うだけなので，それほど計算時間を必要とせず，十分に実用可能である。