中学生のための玉原数学教室　　東京大学
玉原国際セミナーハウス
講義 (2)　小池 祐太
2018 年 10 月 13 日

「素数のはなし」

1. 商と余り

123 を 4 で割ると, 商が 30 で余りが 2 となります. 式で書くと

$$123 = 4 \times 30 + 3$$

と書けるということです. 一般に, a, b を 2 つの整数とするとき, $b > 0$ ならば

$$a = bq + r, \qquad 0 \leq r < b$$

を満たす整数の組 (q, r) がただ一つ存在することが証明できます. このときの q が「a を b で割った**商**」, r が「a を b で割った**余り**」と呼ばれます.

さて, a を b で割った余りが 0 のとき, 「a は b で**割り切れる**」といい, 「a は b の**倍数**である」とか「b は a の**約数**である」ともいいます.

例 1. $123 = 3 \times 41$ であるから, 123 は 3 で割り切れて, 123 は 3 の倍数, 3 は 123 の約数である.

2. 最大公約数

a, b を 2 つの整数とします. a の約数であると同時に b の約数でもあるような正整数のことを, a と b の**公約数**と呼びます. いま, もし a, b がともに 0 でないならば, a と b の公約数は有限個しかないことがわかるので, 最大のものがあります. それを a と b の**最大公約数**と呼びます.

例 2. 8 の約数は 1, 2, 4, 8, 12 の約数は 1, 2, 3, 4, 6, 12 なので, 8 と 12 の公約数は 1, 2, 4 であり, 最大公約数は 4 である.

a と b の最大公約数が 1 であるとき, a と b は**互いに素である**といいます. 例えば, 8 と 9 は互いに素です. 以下で最大公約数を考える場合, 対象となる整数たちは 0 でないとします.

3. ユークリッドの互除法

2 つの整数 a, b の最大公約数を求めるには, 原理的にはそれぞれの整数の約数をすべて数え上げればよいのですが, 整数 a, b が大きい場合にはこれは大変です. 紀元前 (!) のギリシャの数学者ユークリッドは, 約数を数え上げずに最大公約数を求めるための方法を考案しました. この方法の背景には次の原理があります (a は b の倍数でないとします):

「a, b の最大公約数は, a を b で割った余りと b の最大公約数に等しい」

上の原理を使って最大公約数を求める手順を, 具体的な数値例で見てみます.

例 3. 442 と 182 の最大公約数を求める. まず, 442 を 182 で割ってみると

$$442 = 182 \times 2 + 78$$

となるから, 上の原理より 78 と 182 の最大公約数を求めればよい. 次に, 182 を 78 で割ってみると

$$182 = 78 \times 2 + 26$$

となるから, 再び上の原理より 78 と 26 の最大公約数を求めればよい. 引き続き 78 を 26 で割ってみると

$$78 = 26 \times 3$$

となる. すなわち, 26 は 78 の約数であるから, 78 と 26 の最大公約数は 26 である. 以上より, 442 と 182 の最大公約数も 26 である.

上の例のように, 2 つの整数 a, b の最大公約数を求めるには, 大きい方が a, 小さい方が b の場合, a を b で割った余りを r として, r が 0 でなければ b を r で割った余りを r' として, r' が 0 でなければ r を r' で割った余りを r'' として, ..., という操作を余りが 0 になるまで繰り返すと, 最終的に得られた 0 でない余りが答えとなります. この方法は割り算 (除算) を互い違いに繰り返すことで答えを求めるので, **ユークリッドの互除法**と呼ばれています.

ユークリッドの互除法の仕組みから, 特に次のことがわかります.

命題 1. a, b を 2 つの正整数とし, a と b の最大公約数を d とする. このとき,

$$ax + by = d$$

となるような整数 x, y が存在する.

4. 素数

すべての正の整数は 1 および自分自身を約数に持つので, 1 より大きい正の整数は少なくとも 2 つは約数を持つことになります. 特に, 約数をちょうど 2 つ持つような正の整数を**素数**と呼びます. 例えば, 3 の約数は 1 と 3 のみなので素数ですが, 4 は 2 で割り切れるため素数ではありません. 100 以下の素数は全部で 25 個あります (数えてみましょう).

5. 整数の合同

2 で割り切れる正の整数を**偶数**と呼び, それ以外の正整数を**奇数**と呼びます. これは次のように言い換えることができます.

- 偶数とは, 2 で割ったあまりが 0 である正整数のことである.
- 奇数とは, 2 で割ったあまりが 1 である正整数のことである.

このように, 偶数・奇数という分類は, 「整数を 2 で割った余りによる分類」となっています. この考え方を一般化したのが次の「整数の合同」という概念です.

定義 1. m を正の整数とする. 2 つの整数 a と b が m **を法として合同**であるとは, $a - b$ が m で割り切れることをいい, このことを記号

$$a \equiv b \pmod{m}$$

で表す.

「a と b が m を法として合同である」ということは, 「a を m で割った余りと b を m で割った余りが等しい」ということと同じ意味となります.

例 4. $17 - 2 = 15$ は 5 で割り切れるので, $17 \equiv 2 \pmod{5}$ です.

合同式に関しては以下の計算法則が成り立ちます.

命題 2. m を正の整数, a, b, c を整数とするとき, 次のことが成り立つ.

反射律: $a \equiv a \pmod{m}$.
対称律: $a \equiv b \pmod{m}$ ならば, $b \equiv a \pmod{m}$.
推移律: $a \equiv b \pmod{m}$ かつ $b \equiv c \pmod{m}$ ならば, $a \equiv c \pmod{m}$.

命題 3. m を正の整数, a, b, c, d を整数とし, $a \equiv b \pmod{m}$ かつ $c \equiv d \pmod{m}$ が成り立つとする. このとき, 次のことが成り立つ.

(a) $a + c \equiv b + d \pmod{m}$.
(b) $ac \equiv bd \pmod{m}$.
(c) n を正の整数とすると, $a^n \equiv b^n \pmod{m}$.

命題 4. p, q を正の整数, a, b を整数とする. このとき, $a \equiv b \pmod{p}$, $a \equiv b \pmod{q}$ かつ p と q が互いに素ならば, $a \equiv b \pmod{pq}$ が成り立つ.

9 去法と 11 去法. ある正整数 a が別の整数 m で割り切れるかどうかを判定するのは, a の桁数が大きい場合は暗算では難しいことが多いですが, 特殊な m については判定を簡略化することができます. 例えば $m = 2$ の場合, すなわち a が偶数か奇数かを判定する場合は, a の 1 桁目が偶数か奇数かで判定することができます. なぜこのようにして良いかということは合同式を使うとよくわかります. いま, a は N 桁の正整数であるとして, a の i 桁目を a_i で表すことにしましょう. 従って

$$a = a_1 + 10a_2 + \cdots + 10^{N-1}a_N$$

と書き表すことができます. 例えば, $a = 12345678$ ならば, $N = 8$ であり,

$$a = 8 + 10 \cdot 7 + 10^2 \cdot 6 + 10^3 \cdot 5 + 10^4 \cdot 4 + 10^5 \cdot 3 + 10^6 \cdot 2 + 10^7 \cdot 1$$

となります. ここで, すべての正整数 n について, 10^n は 2 で割り切れますから,

$$10^n \equiv 0 \pmod{2}$$

が成り立ちます. 従って, 命題 3(a) を繰り返し使うと,

$$\begin{aligned} a &\equiv a_1 + 10a_2 + \cdots + 10^{N-1}a_N \pmod{2} \\ &\equiv a_1 + 10a_2 + \cdots + 10^{N-2}a_{N-1} \pmod{2} \\ &\cdots \\ &\equiv a_1 \pmod{2} \end{aligned}$$

となります. よって, a が 2 で割り切れることと a_1 が 2 で割り切れることは同じ意味となるため, a が偶数か奇数かは a_1 が偶数か奇数かで判断して良いこととなります. 例えば, 12345678 の 1 桁目 8 は偶数ですから, 12345678 は偶数ということになります.

この考え方を応用すると, 次のような判定法が得られます.

- a が 3 で割り切れるためには, a のすべての桁を足したもの $a_1 + a_2 + \cdots + a_N$ が 3 で割り切れればよいです. 実際, $10 \equiv 1 \pmod{3}$ なので, 命題 3(c) よりすべての正整数 n について

$$10^n \equiv 1 \pmod{3}$$

が成り立ちます. 従って命題3(b)より $10^n a_n \equiv a_n \pmod 3$ がわかりますので, 命題3(a)を繰り返し使って,

$$a \equiv a_1 + a_2 + \cdots + a_N \pmod 3$$

が得られます. 例えば, $1+2+3+4+5+6+7+8 = 36$ は3で割り切れるので, 12345678は3で割り切れます.

- 上の例と同様の考え方で, a が9で割り切れるためには, a のすべての桁を足したもの $a_1 + a_2 + \cdots + a_N$ が9で割り切れればよいこともわかります. 従って12345678は9でも割り切れます.
- a が11で割り切れるためには, $a_1 - a_2 + a_3 - a_4 + \cdots + (-1)^{N-1} a_N$ が11で割り切れればよいです. これは, $10 \equiv -1 \pmod{11}$ であることから, 上と同様の議論によって

  $$a \equiv a_1 - a_2 + a_3 - a_4 + \cdots + (-1)^{N-1} a_N \pmod{11}$$

  がわかるからです. 例えば, $8-7+6-5+4-3+2-1 = 4$ は11で割り切れないので, 12345678は11で割り切れません.

6. 素数を使った暗号の作成

ウェブ上で何らかのサービスを利用するときは, 不正利用を防ぐためにパスワードを入力するのが普通です. このとき, 入力されたパスワードが送信中に盗聴されても大丈夫なように, 暗号化してから送信するのがセキュリティ上望ましいです. このためには, (通常不特定多数の) 利用者側がパスワードを暗号化する方法を知っている必要がありますが, 一方で暗号文を復元する方法はサービスを提供する会社のみが知っている, という状況を作る必要があります (従って, 「タヌキの書いた手紙」のような暗号化ではダメだということです). このような状況を作るために役立つ暗号として**公開鍵暗号**というものがあります. 公開鍵暗号では, 暗号化の方法は不特定多数の人々に公開されていますが, 暗号化の方法を知っているだけでは暗号文を復元するのが困難であるように暗号が設計されています.

有名な公開鍵暗号の1つに **RSA 暗号**と呼ばれるものがありますが, これは素数の性質を使って設計された暗号ですので, ここでは簡単にその仕組みを見てみることにします.

まず, 受信者側は以下の準備をします.

(1) 2つの異なる素数 p, q を選び, $n = pq$ と $f = (p-1)(q-1)$ を計算する.
(2) f と互いに素な正整数 e を1つ選び, $de \equiv 1 \pmod f$ となるような正整数 d を1つ求める (そのような d は命題1から必ず存在する).
(3) p, q, d は秘密にして, n, e を公開する.

発信者側が n より小さい正整数 a を暗号化して送信し, 受信者側がそれを復元するには以下の手順を実行します.

暗号化: a^e を n で割った余り x を発信者側に送信する.
復元: x^d を n で割った余りが元の整数 a となる.

例 5. 例として最も簡単な p, q の選び方 $p = 3, q = 5$ を考えてみる. この場合 $n = 3 \cdot 5 = 15$, $f = 2 \cdot 4 = 8$ である. 従って例えば $e = d = 3$ と取ることができる. この場合, 公開するのは $n = 15$ と $e = 3$ である.

さて, 1013というパスワードが送信されたとき, これを2桁ずつRSA暗号によって暗号化してみると,

- $10^3 = 1000$ を 15 で割った余りは 10
- $13^3 = 2197$ を 15 で割った余りは 7

であるから, 暗号文は 10-7 となる. いま,

- $10^3 = 1000$ を 15 で割った余りは 10
- $7^3 = 343$ を 15 で割った余りは 13

であるから, 復元の手順によって元のパスワード 1013 が復元できる.

RSA 暗号の安全性は, 「非常に大きい整数を 2 つの素数の積に分解するのは非常に時間がかかる」という事実に依拠しています. 逆にいうと, RSA 暗号の安全性を保証するには, n が非常に大きな値となるように p, q を選ぶ必要があります (従って上の例のような p, q の選び方は実際には使われません). 他にも, RSA 暗号の安全性を保証するには p, q の選び方について気をつけるべき点がいくつかあることが知られていますが, 詳細は割愛します.

最後の手順で暗号が復元できる原理. 最後の手順で暗号が復元できるのは,

$$x^d \equiv a \pmod{n} \tag{1}$$

が成り立つからです. では, なぜこの式は成り立つのでしょうか. その証明には次の定理を使います.

定理 1 (フェルマーの小定理)**.** p を素数とし, a を p と互いに素な整数とする. このとき,

$$a^{p-1} \equiv 1 \pmod{p}$$

が成り立つ.

(1) **式が成り立つことの証明.** 暗号化の手順から $a^e \equiv x \pmod{n}$ だから, 命題 3(c) より $(a^e)^d \equiv x^d \pmod{n}$, 従って $(a^e)^d \equiv a \pmod{n}$ を示せば推移律より示すべき結論が得られる. $n = pq$ であり, p, q は互いに素であるから, 命題 4 より次の 2 つの式を示せば十分である:

$$(a^e)^d \equiv a \pmod{p}, \tag{2}$$

$$(a^e)^d \equiv a \pmod{q} \tag{3}$$

両式とも同様の議論で証明できるので, (2) だけ証明する. まず, もし a が p で割り切れるならば, $(a^e)^d$ も p で割り切れるので, (2) は成り立つ. 従って a が p で割り切れない場合だけ考えれば十分である. この場合, p が素数であることから, a と p は互いに素でなければならない. フェルマーの小定理より $a^{p-1} \equiv 1 \pmod{p}$ が成り立つので, 命題 3(c) より $(a^{p-1})^{q-1} \equiv 1 \pmod{p}$ が成り立つ. 指数法則より $a^f = a^{(p-1)(q-1)} = (a^{p-1})^{q-1}$ であるから, これは $a^f \equiv 1 \pmod{p}$ を意味する. さて, $de \equiv 1 \pmod{f}$ であったから, ある整数 q が存在して $de = fq + 1$ と書ける. このとき, 再び指数法則より $(a^e)^d = a^{de} = a^{fq+1} = (a^f)^q \cdot a$ が成り立つ. いま, 再び命題 3(c) を適用すると $(a^f)^q \equiv 1 \pmod{p}$ を得るので, 命題 3(b) より $(a^f)^q \cdot a \equiv a \pmod{p}$ が成り立つ. すなわち $(a^e)^d \equiv a \pmod{p}$ であるから (2) が証明された. □