Algorithms – CMSC-27200
Basic algorithms in Number Theory: Euclid's algorithm and multiplicative inverse

Instructor: László Babai

Last updated at 12:40pm on 10-19-2022.

$\mathbb{Z}$ denotes the set of integers. All variables in this note range over $\mathbb{Z}$.

1 Divisibility, definition of gcd

Definition 1.1 (divisibility)**.** We say that a divides b (written as $a \mid b$) if $(\exists x)(ax = b)$.

Exercise 1.2. Prove: (a) d divides all numbers $\iff$ $d = \pm 1$.
(b) All numbers[1] divide z $\iff$ $z = 0$.

Exercise 1.3. Prove: (a) If $d \mid a$ and $d \mid b$ then $d \mid a \pm b$ (additivity of divisibility);
(b) If $a \mid b \mid c$ then $a \mid c$ (transitivity of divisibility).
In each case, state, what basic property of multiplication you are using (associativity, distributivity, etc.).

Let $\mathrm{Div}(a)$ denote the set of divisors of the number a. For instance,

$$\mathrm{Div}(-6) = \{\pm 1, \pm 2, \pm 3, \pm 6\}. \tag{1}$$

Exercise 1.4. Prove: (a) $\mathrm{Div}(a) = \mathrm{Div}(-a)$. (b) $\mathrm{Div}(0) = \mathbb{Z}$
(c) If $a \neq 0$ then $\mathrm{Div}(a)$ is a finite set.

Exercise 1.5. Prove that the following three statements are equivalent:
(a) $a \mid b$ and $b \mid a$
(b) $\mathrm{Div}(a) = \mathrm{Div}(b)$
(c) $a = \pm b$

Notation 1.6.

$$\mathrm{Div}(a, b) = \mathrm{Div}(a) \cap \mathrm{Div}(b)\,. \tag{2}$$

So $\mathrm{Div}(a, b)$ is the set of **common divisors** of a and b.

Definition 1.7 (gcd)**.** We say that d is a greatest common divisor of a and b if

(i) $d \mid a$ and $d \mid b$ (: d is a common divisor of a and b :)

(ii) $(\forall e)($ if $e \mid a$ and $e \mid b$ then $e \mid d)$
(: d is divisible by all common divisors of a and b :)

So d is "greatest" in the sense of divisibility. — We rephrase the definition. The following exercise is central to our discussion.

[1] Part (b) of Ex. 1.2 includes the fact that $0 \mid 0$. The reader may be puzzled: did we overrule the prohibition agains division by zero? No, we did not: the definition of divisibility (Def. 1.1) does not involve division, only multiplication. So $0 \mid 0$ because there exists x such that $0 \cdot x = 0$ (for instance, $x = 17$ is a solution).

Exercise 1.8. Prove: d is a greatest common divisor of a and b if and only if

$$\mathrm{Div}(a,b) = \mathrm{Div}(d), \tag{3}$$

i. e., the common divisors of a and b are precisely the divisors of d.

Note that if the number d satisfies either definition then so does $-d$. The next exercise asserts that there are no other greatest common divisors.

Exercise 1.9. If d is a greatest common divisor of a and b then the greatest common divisors of a and b are d and $-d$.

To make the gcd notation unique, we additionally require $\gcd(a,b) \geq 0$.

Convention 1.10. If d is a greatest common divisor of a and b then we write $\gcd(a,b) = |d|$.

So for example $\gcd(-30,-30) = 30$.

Note that by Ex. 1.9, $\gcd(a,b)$ is **unique**, assuming it exists.

We defined $\gcd(a,b)$ by a **wish-list**; we need to prove that such a d always **exists**. The proof will be algorithmic: it will not only prove the existence of such d but also give an efficient way to calculate d.

Theorem 1.11. *Every pair of numbers has a greatest common divisor.*

First we settle some special cases.

Exercise 1.12. $\mathrm{Div}(a,0) = \mathrm{Div}(a)$. Therefore $\gcd(a,0)$ exists and is equal to $|a|$. (Note in particular that $\gcd(0,0) = 0$.)

Exercise 1.13. $\mathrm{Div}(a,a) = \mathrm{Div}(a)$. Therefore $\gcd(a,a)$ exists and is equal to $|a|$.

We also note that

Exercise 1.14. (a) $\mathrm{Div}(a) = \mathrm{Div}(|a|)$ (b) $\mathrm{Div}(a,b) = \mathrm{Div}(|a|,|b|)$ (c) $\mathrm{Div}(a,b) = \mathrm{Div}(b,a)$.

The key lemma to the proof of Theorem 1.11 is the following.

Exercise 1.15 (Euclid's gcd Lemma, modern version)**.**

$$\mathrm{Div}(a,b) = \mathrm{Div}(a-b,b)\,.$$

This lemma is not to be confused with "Euclid's Lemma" which says that if p is a prime number and $p \mid ab$ then $p \mid a$ or $p \mid b$.

It follows from Euclid's gcd Lemma that

Exercise 1.16. For all integers k we have $\mathrm{Div}(a,b) = \mathrm{Div}(a-kb,b)$.

(First prove this for $k \geq 0$ by inductions on k and then infer from this that the result also holds for all negative k.)

To make the algorithm efficient, we need to review the Division Theorem.

Exercise 1.17 (Division Theorem)**.**

$$(\forall a,b)\,(\text{ if } b \neq 0 \text{ then } (\exists q,r)(a = bq + r \text{ and } 0 \leq r < |b|))$$

(Prove by induction on $|a|$.) The quantity q is the "integer quotient of a divided by b" and r is the "least non-negative remainder of a modulo b."

Notation 1.18. We write $r = (a \bmod b)$ to denote the least non-negative remainder of a modulo b.

For instance, $2 = (30 \bmod 7) = (30 \bmod -7)$ and $5 = (-30 \bmod 7)$. This is the LaTeX code[2] for such expressions:

```
$5 = (-30 \bmod{7})$
```

2 Proof of existence of gcd: Euclid's algorithm

To prove Theorem 1.11, we pick two numbers, a, b, of which we wish to compute the gcd. By Ex. 1.14 we may assume $a \geq b \geq 0$.

Procedure: Euclid's algorithm
Input: integers $a \geq b \geq 0$
Output: $\gcd(a, b)$

Pseudocode A.

```
Initialize: $A := a$, $B := b$
while $B \geq 1$ do
    $R := (A \bmod B)$        [Division Theorem]
    $A := B$, $B := R$
end(while)
return $A$
```

Exercise 2.1. Prove that the algorithm terminates in a finite number of steps.

The **correctness** of the algorithm follows from the following *loop invariant:*

$$\mathrm{Div}(A, B) = \mathrm{Div}(a, b)\,. \tag{4}$$

Exercise 2.2. Prove that this is indeed a loop invariant.

This means the following: If Eq. (4) holds when the algorithm enters the **while** loop then Eq. (4) also holds when the algorithm exits the **while** loop. Given this loop invariant, we conclude that when looping terminates (i. e., when $B = 0$) we have $\mathrm{Div}(a, b) = \mathrm{Div}(A, 0) = \mathrm{Div}(A)$, so the returned value A qualifies as $\gcd(a, b)$ by Ex. 1.8. This completes the proof of correctness of Euclid's algorithm and thereby the proof of Theorem 1.11 (that gcd exists). □

The **efficiency** of the algorithm follows from the following observation:

[2] The pronunciation of the term "LaTeX" is not "latex" (the source of natural rubber or rubbery synthetic materials). It derives from TeX, the typesetting system created by Donald Knuth, on which LaTeX is built. As professor Knuth explains in his TeXBook, the three letters are not the Roman T, E, X but the Greek τ, ϵ, χ and should be pronounced like the first syllable in words like "technical" or "technology."

Exercise 2.3. (a) Let B_i be the value of B produced after the i-th iteration of the **while** loop, starting with $B_0 = b$. Prove: for all i we have $B_{i+2} \leq B_i/2$.

(b) Suppose b has n digits in binary. Then Euclid's algorithm terminates in $\leq 2n$ rounds. (A *round* is one execution of the **while** loop.)

Exercise 2.4 (Division algorithm)**.** If A and B are positive integers with $\leq n$ binary digits then $A \bmod B$ can be computed in $O(n^2)$ bit operations.

Therefore the total number of bit-operations is $O(n^3)$, so this is a *polynomial-time algorithm.* (Good job, Euclid!)

Exercise 2.5. (a) Modify the Division Theorem so r satisfies $-|b|/2 < r \leq |b|/2$ (remainder with least absolute value). (b) Show that if in Euclid's algorithm we use remainders with least absolute value, the process terminates in $\leq n$ rounds (where, as before, n is the number of binary digits of b).

Exercise 2.6 (Euclid's gcd Lemma, original version)**.**

$$\gcd(a, b) = \gcd(a - b, b) .$$

3 Recursive implementation

We now give an alternative, recursive implementation of Euclid's algorithm. The non-recursive code is preferred.

Pseudocode B: recursive.

```
procedure gcd($a, b$)    ($a \geq b \geq 0$)
    if $b = 0$ then return $a$
    else $r := (a \bmod b)$        [Division Theorem]
        return gcd($b, r$)
```

Exercise 3.1. Analyse this version of the algorithm.

4 Gcd as linear combination

Definition 4.1. A *linear combination* of the numbers a, b is a number of the form $au + bv$. (As everywhere in this note, all numbers mentioned are integers.)

A fundamental result about the gcd says that it can be written as a linear combination:

Theorem 4.2 (Bézout's Lemma)**.** $(\forall a, b)(\exists u, v)(\gcd(a, b) = au + bv)$

For instance, $\gcd(72, 52) = 4 = 72 \cdot (-5) + 52 \cdot 7$

Exercise 4.3. Prove Theorem 4.2 by induction on $|a| + |b|$.

Hint: use Euclid's gcd Lemma.

Exercise 4.4. Prove: if d is a common divisor of a and b and d is a linear combination of a and b then d is a greatest common divisor of a and b (i. e., $|d| = \gcd(a, b)$).

5 Congruence, modular arithmetic

Definition 5.1 (Congruence)**.** We say that a is **congruent**[3] to b modulo m if $m \mid a - b$. Notation: $a \equiv b \pmod{m}$.

The notation in LaTeX:

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

Exercise 5.2. Prove that congruence modulo m is an equivalence relation on $\mathbb{Z}$.

The equivalence classes are called *residue classes mod m.* If $m \neq 0$ then there are exactly $|m|$ residue classes mod m. NOTE: a and b **belong to the same residue class modulo** m if and only if $a \equiv b \pmod{m}$.

Exercise 5.3. Prove: if $a \equiv x \pmod{m}$ and $b \equiv y \pmod{m}$ then
$a \pm b \equiv x \pm y \pmod{m}$ and $ab \equiv xy \pmod{m}$.

Exercise 5.4. Prove: If $a \equiv b \pmod{m}$ then $\gcd(a, m) = \gcd(b, m)$. In other words, all members of a residue class mod m have the same gcd with the modulus.

6 Multiplicative inverse

Definition 6.1. We say that x is a multiplicative inverse of a modulo m if $ax \equiv 1 \pmod{m}$.

Exercise 6.2. Suppose x is a multiplicative inverse of a modulo m. Then y is a multiplicative inverse of a modulo m if and only if $y \equiv x \pmod{m}$.

So the multiplicative inverse is unique modulo m. We denote the least positive multiplicative inverse of a by $(a^{-1} \bmod m)$.
So for instance $(3^{-1} \bmod 17) = 6$.

Theorem 6.3. *The number $\underline{a}$ has a multiplicative inverse mod m* $\iff$ $\gcd(a, m) = 1$.

The $\Rightarrow$ direction of the proof is easy.

Exercise 6.4. Prove: if $a \equiv b \pmod{m}$ and $d \mid m$ then $a \equiv b \pmod{d}$.

Exercise 6.5. Prove the $\Rightarrow$ direction of Theorem 6.3.

Hint: let $d = \gcd(a, m)$. Suppose x is a multiplicative inverse of a modulo m. So $ax \equiv 1 \pmod{d}$. But $ax \equiv 0 \pmod{d}$. So $1 \equiv 0 \pmod{d}$.

Exercise 6.6. Prove the $\Leftarrow$ direction of Theorem 6.3.

Hint. Apply Bézout's lemma (Theorem 4.2) to a and m.

[3] The text uses the term "a is *equivalent* to b modulo m." This is not the commonly used term in this context. Please use "congruent."

7 Computing the multiplicative inverse

Next we shall learn how to use Euclid's algorithm to efficiently compute the multiplicative inverse.

We illustrate this on the example of computing $x = (13^{-1} \bmod 18)$.

$\gcd(18, 13) = 1$, so x exists. We can establish that $\gcd(18, 13) = 1$ through the following sequence of remainders: $18, 13, 5, 3, 2, 1, 0$ (namely, $18 = 13 \cdot 1 + 5$, $13 = 5 \cdot 2 + 3$, $5 = 3 \cdot 1 + 2$, $3 = 2 \cdot 1 + 1$, $2 = 1 \cdot 2 + 0$).

Consider now the following congruences. Both of these are true for $x = (13^{-1} \bmod 18)$.

$$\begin{aligned} 18x &\equiv \;\; 0 \pmod{18} \\ 13x &\equiv \;\; 1 \pmod{18} \end{aligned} \tag{5}$$

Subtracting the second from the first, we obtain

$$5x \equiv \;\; -1 \pmod{18} \tag{6}$$

Subtracting twice this congruence from the preceding one, we obtain

$$3x \equiv \;\; 3 \pmod{18} \tag{7}$$

Subtracting this congruence from the preceding one, we obtain

$$2x \equiv \;\; -4 \pmod{18} \tag{8}$$

Subtracting this congruence from the preceding one, we obtain

$$x \equiv \;\; 7 \pmod{18} \tag{9}$$

So the solutions are those values of x that are congruent to 7 modulo 18. In particular, $(13^{-1} \bmod 18) = 7$. (Verify!)

REMARK. Let us review the **logic** of this argument. We started from the pair of congruences (5) and from these we deduced congruence (8). Does this mean that x is a multiplicative inverse of 13 modulo 18 if and only if $x \equiv 7 \pmod{18}$? NO. What it means is this: IF x is a multiplicative inverse of 13 modulo 18 THEN $x \equiv 7 \pmod{18}$. This argument does NOT in itself prove that $x = 7$ or $x = 25$ or $x = -11$ etc. are solutions. What we proved is that ONLY these numbers (those that are congruent to 7 modulo 18) can be solutions. But Theorem 6.3 guarantees that a solution EXISTS. So we knew from the beginning that the number x we are looking for exists. We concluded that x must belong to the residue class of numbers that are $\equiv 7 \pmod{18}$. Does that mean that for instance $x = 7$ is a solution? This is now guaranteed by Exercise 6.2 which says in particular that the multiplicative inverses form an entire residue class modulo 18. So if there is a number in the residue class $x \equiv 7 \pmod{18}$ that is a multiplicative inverse of 13 modulo 18 then all numbers in this residue class (including 7, 25, -11, etc.) are. So it is a combination of the calculations above with Theorem 6.3 and Exercise 6.2 that guarantee that every number $\equiv 7 \pmod{18}$ is a multiplicative inverse of 13 modulo 18, and no number outside this residue class is. We found all solutions. — Was it necessary to *verify* our solution at the end? Not if we are confident that we did not make an arithmetic error. The residue class $x \equiv 7 \pmod{18}$ is guaranteed to consist precisely of the multiplicative inverses of 13 modulo 18. The only purpose of the verification is to check we did not make an error in the calculations.

The next exercise asks you to turn the scheme we used above to calculate $(13^{-1} \bmod 18)$ into an algorithm.

Exercise 7.1. Suppose $\gcd(a, m) = 1$. Design an efficient algorithm that finds $(a^{-1} \bmod m)$. Write your algorithm in elegant pseudocode. Do not use recursion.

Hint: your code should be a modification of Pseudocode A. Your algorithm should be as efficient as Euclid's (within a constant factor).

Exercise 7.2. Given a and b, write $\gcd(a, b)$ as a linear combination of a and b. Design an efficient algorithm to do so. Explain your algorithm in concise, unambiguous English, no pseudocode required.

Hint. Here is an example: Let $a = 72$ and $b = 52$. Then $\gcd(a, b) = 4$. We need to find u, v such that $4 = 72u + 52v$. Divide each side of this equation by 4; we get $1 = 18u + 13v$. So v must be a multiplicative inverse of 13 modulo 18. (Why?) So take $v = 7$ (computed above), and find u. (Compare with the example given after Theorem 4.2.)