Modern Algebra I: The Euclidean algorithm

As promised in the lecture, we describe a computationally efficient method for finding the gcd of two positive integers a and b, which at the same time shows how to write the gcd as a linear combination of a and b.

Begin with a, b. Write $a = bq_1 + r_1$, with integers q_1 and r_1, $0 \leq r_1 < b$. Note that $r_1 = a + b(-q_1)$ is a linear combination of a and b. If $r_1 = 0$, stop, otherwise repeat this process with b and r_1 instead of a and b, so that $b = r_1 q_2 + r_2$, with $0 \leq r_2 < r_1$, and note that $r_2 = b - r_1 q_2 = b - aq_2 + bq_1 q_2$ is still a linear combination of a and b. If $r_2 = 0$, stop, otherwise repeat again with r_1 and r_2 instead of b and r_1, so that $r_1 = r_2 q_3 + r_3$, with $0 \leq r_3 < r_2$. We can continue in this way to find $r_1 > r_2 > r_3 > \cdots > r_k \geq 0$, with $r_{k-1} = r_k q_{k+1} + r_{k+1}$. Since the sequence of the r_i decreases, and they are all nonnegative integers, eventually this procedure must stop with an r_n such that $r_{n+1} = 0$, and hence $r_{n-1} = r_n q_{n+1}$. The procedure looks as follows:

$$
\begin{aligned}
a &= bq_1 + r_1 \\
b &= r_1 q_2 + r_2 \\
r_1 &= r_2 q_3 + r_3 \\
&\vdots \\
r_{n-2} &= r_{n-1} q_n + r_n \\
r_{n-1} &= r_n q_{n+1}.
\end{aligned}
$$

We claim that r_n is the gcd of a and b. In fact, we shall show:

(i) r_n divides both a and b;

(ii) r_n is a linear combination of a and b.

(i) Since $r_n | r_{n-1}$, the equation $r_{n-2} = r_{n-1} q_n + r_n$ implies that $r_n | r_{n-2}$, and then working backwards from the equation $r_{k-1} = r_k q_{k+1} + r_{k+1}$, we see (with reverse induction) that $r_n | r_{k-1}$ for all $k < n$. The fact that $b = r_1 q_2 + r_2$ and that r_n divides r_1 and r_2 implies that r_n divides b, and then the equation $a = bq_1 + r_1$ implies that r_n divides a, too.

(ii) Working the other way, we have seen that r_1 and r_2 are linear combinations of a and b. By induction, if r_{k-1} and r_k are linear combinations of a and b, then the equation $r_{k-1} = r_k q_{k+1} + r_{k+1}$ implies that $r_{k+1} = r_{k-1} - r_k q_{k+1}$ is also a linear combination of a and b (because as we saw in class the set of all linear combinations of a and b is a subgroup of $\mathbb{Z}$ and thus is closed under addition, subtraction, and multiplication by an integer). Thus r_n is a linear combination of a and b as well. But we have seen that if a linear combination of a and b divides a and b and is positive, then it is equal to the gcd of a and b. So r_n is the gcd of a and b.

The algorithm is easier to carry out than it is to explain! For example, to find the gcd of 34 and 38, we have

$$\begin{aligned} 38 &= 34(1) + 4 \\ 34 &= 4(8) + 2 \\ 4 &= 2(2). \end{aligned}$$

This says that $2 = \gcd(34, 38)$ and that $2 = 34 - 4(8) = 34 - (38 - 34)(8) = 9(34) + (-8)(38)$.

It is often more efficient to choose q_{k+1} and r_{k+1} so that $r_{k-1} = r_k q_{k+1} \pm r_{k+1}$, with $r_{k+1} < r_k$ and the sign chosen so that r_{k+1} is as small as possible. In other words, we allow negative remainders of the form $-r_k$ with the goal of minimizing the absolute value of the remainder. For example, to find the gcd of 7 and 34, we could write

$$\begin{aligned} 34 &= 7(4) + 6 \\ 7 &= 6(1) + 1, \end{aligned}$$

to see that the gcd is 1 and that $1 = 7 - 6 = 7 - (34 - 4(7)) = -34 + 5(7)$, or we could see directly that

$$34 = 7(5) - 1.$$

A more complicated example is the following, to find the gcd of 1367 and 298:

$$\begin{aligned} 1367 &= (298)(5) - 123 \\ 298 &= 123(2) + 52 \\ 123 &= 52(2) + 19 \\ 52 &= 19(3) - 5 \\ 19 &= 5(4) - 1. \end{aligned}$$

Thus the gcd is 1, and a little patience shows that

$$\begin{aligned} 1 &= 5(4) - 19 = 11(19) - 4(52) = 11(123) - 26(52) = \\ &= (63)(123) - (26)(298) = (-63)(1367) + (289)(298). \end{aligned}$$