Integer Factorization Using Digit Strings

Written by Clinton Morrison

Integer factorization is an interesting problem that has been studied in great depth. The digits of integers follow interesting patterns related to their divisibility. Divisors of certain integers can be inferred from only looking at their digits. While it is only a relatively small class of integers for which this is useful, it is still interesting to consider. This short paper explores a few simple properties that come from examination of digits in numbers and how they relate to divisibility.

Notation

Some notation is required to properly examine the digits of an integer. First, let d_n give the number of digits in the number n. Let a_i be the i^{th} digit from the right of the number. Thus $a_0$is the rightmost digit, and a_{d_n-1} is the leftmost digit. For example, if $n = 1234$ then $d_n = 4, a_0 = 4, a_1 = 3, a_2 = 2, a_3 = 1$. Finally, use of brackets around a number donates a number is given in digit form. For example if $a = 1, n = [aaa] = 111$

In the general case we can write *n* as follows:

$$n = \left[a_{d_n-1}\, a_{d_n-1} \ldots a_2\, a_1\, a_0\right] = a_0 + 10a_1 + \ldots + 10^{d_n-1}a_{d_n-1} = \sum 10^i a_i,\ a \in \mathbb{Z},\ 0 \le a_i \le 9$$

Special Forms

It is obvious that if *n* is composite then the above expression must be factorable. Certain specific cases of digit values can be considered. For example the case where a number consists of a string a single digit, i.e. $n = [aaaa]$. Then we can write:

$$n = a + 10a + 100a + 1000a = a(1 + 10 + 100 + 1000) = 1111a$$

Thus numbers of the form $n = [aaaa]$ are all divisible by 1111, as well as a. Similar arguments can be made for numbers of other forms. For example where $a_2 = 2a_0 = 2a_1$, i.e. $n = 633$

Divisibility given certain rightmost digits

It follows from a similar argument that all numbers ending in an even number are divisible by 2, and numbers ending in "0" and "5" and divisible by 10 and 5 respectively. We can write n as

$$n = a_0 + 10a_1 + 100a_2 + \ldots + 10^{d(n)}a_{d(n)} = a_0 + 2(5)a_1 + 2^2(5^2)a_2 + \ldots + 2^{d_n}(5^{d_n})a_{d_n}$$

Clearly, if $2|a_0$ (2 divides a_0) or if $a_0 = 0$ then a 2 could be factored out of the expression and thus $2|n$. If $a_0 = 5$ or $a_0 = 0$ then it is also clear that a 5 could be factored out, and so $5|n$. If $a_0 = 0$ then $10|n$.

Number base and factorability

So far only numbers in base 10 have been discussed. It is interesting to notice that the special cases where a_0 allows us to infer the divisibility of n occur only where $a_0|10$. This is not coincidental. Indeed, relationship between the rightmost digit and the factorability of n is determined by the base that the number is expressed in. We know the number can be written as a sum where every term excluding a_0 is necessarily divisible by 10. If $a_0|10$ then $a_0|n$. Consider more generally a number n written in a base B. We can write n as

$$n = a_0 + Ba_1 + \ldots + B^{d_n-1}a_{d_n-1} = \sum B^i a_i, \qquad a \in \mathbb{Z}, \qquad 0 \le a_i \le B-1$$

If B is divisible by a_0 then a_0 can be factored out of every term in the above sum. It follows that

$$a_0|B \Rightarrow a_0|n$$

$$a_0 = 0 \Rightarrow B|n$$

It is interesting to note that this implies that numbers in prime bases cannot be easily factored by considering only the last digit. The only a_0 for which $a_0|B$ would be where $a_0 = 0$, given that $B \in \mathbb{P}$. Conversely, the last digit would yield most information for numbers in bases with many unique prime factors.

GCD of digits and divisibility

Another interesting point relates to the greatest common divisor of the digits, $gcd(a_1, a_2, a_3, \ldots, a_{d_n-1})$. Specifically, there is a nontrivial greatest common if there exists a number b where $b|a_0$, $b|a_1$, $b|a_2, \ldots, b|a_{d_n-1}$ and $b > 1$, $b \in \mathbb{Z}$. If this is the case then it can be said that:

$$n = b\left(\frac{a_0}{b} + \frac{10a_1}{b} + \frac{100a_2}{b} + \ldots + \frac{10^{d(n)}a_{d(n)}}{b}\right)$$

Then we know $b|n$. This is useful as long as $b > 1$.

Generalization to digit strings of arbitrary size

The ideas discussed above pertained only to single digits within the number. It is possible and useful to generalize the ideas to analyze strings of digits of arbitrary size. Let $a_{j\ldots i} = [a_j a_{j-1} \ldots\ a_{i+1} a_i]$. $a_{j\ldots i}$ and $a_{k,j,i} = [a_k a_j a_i]$. For example, if $n = 223344,\ a_{3\ldots 0} = 2334$, $a_{5,4} = 22$, $a_{4,3,2} = 233$, etc.

Then we can write *n* as

$$n = \sum B^i a_{j\ldots i}\,, a \in \mathbb{Z}, a \geq 0$$

For example if we break *n* into 2 digit numbers we get

$$n = a_{1,0} + B^2 a_{3,2} + B^4 a_{5,4} + \cdots + B^{d_n-2} a_{\,d_n-1,d_n-2}$$

The length of the digit string is trivial and need not be constant. Unfortunately $a_{1,0}$ alone does not allow us to infer anything more about the factors of n than a_0 does. This is because the other terms are only guaranteed to be divisible by B, and so a_0 alone must be divisible by B. However, the result regarding GCD is still relevant. If $b = gcd\left(a_{i..0,}\ldots, a_{d(n)..j}\right)$ then $b|n$.

Summary

While the results presented here are not novel and generally not useful, they are interesting. The notation and general process explained here can be applied to quickly infer factors of numbers of certain special forms. Numbers whose base is divisible by their rightmost digit are divisible by that digit. If the rightmost digit of a number is 0 then that number is divisible by the base it is expressed in. Additionally, numbers are necessarily divisible by the greatest common divisor of the digits which make up that number. Such numbers are also divisible by the greatest common divisor of strings of digits within the number, given that all the digits are represented by this set of strings.