ON THE NUMBER OF NUMBERS WITH A GIVEN DIGIT SUM

Jon Perry
51 High Street, Belbroughton , West Midlands, DY9 9ST, England

Abstract We consider the sum of digits function which maps an integer to the sum of it's digits, for example 142 is mapped to $1 + 4 + 2 = 7$. This papers examines the question of how many other integers are mapped to a given digit in the range 1 to 10^z.

§1. Introduction

To begin with, we need a sum of digits function [1]. The code in this paper has been written in Pari/GP [2].

x is used to determine the number of digits of n, and d is used to store the cumulative sum of the digits, which are extracted by considering the last digit, and removing it, and then divide by 10.

To test the function, and to see the values to $n = 100$:
for $(n = 1, 100)$,print1$(", "sd(n))$

$1, 2, 3, 4, 5, 6, 7, 8, 9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19$

This is $A007953$ at the Online Encyclopaedia of Integer Sequences [3].

The function has obvious patterns. We will only look at digit sums to 9, as the theory for higher digit sums is too complex for this paper.

§2. Partition Theory.

The sum of digits function is closely related to the theory of partitions and compositions.

A partition of n is a sum $d_1 + d_2 + \ldots + d_k$, for some k less than or equal to n such that the sum is n, and the d's are ordered from largest to smallest.

For example, the partitions of 4 are $4, 3+1, 2+2, 2+1+1$ and $1+1+1+1$.

A composition is the same, except for the order of the d's doesn't matter, so the compositions of 4 are as with the partitions, but also including $1+3$, $1+2+1$ and $1 + 1 + 2$.

With the sum of digits we need a new concept, as we are dealing with compositions that can be 'stretched' using zeroes in-between the digits.

We also need to add the constraint that any d can be no larger than 9.

The picture is further complicated by the fact that as we consider larger integers, we have more stretched compositions associated with any given composition.

For example, if we consider the composition $1+3$, this is directly associated with 13.

In three digits, we may have 103 and 130. In four digits we can have $1003, 1030$ and 1300.

Easy enough, but try and calculate the number of stretched compositions of 121 over eight digits!

§3. Length of Compositions

To do this, we consider a composition of n, and we define the number of individual digits in a sum as its length. $1 + 2 + 1$ has length 3, etc...

We know that there are 2^{n-1} compositions of n, but determination of the length of each composition is more complicated.

The following tables examine this up to $n = 4$:

The sum of the lengths is $1, 3, 8, 20$, which is [4], and has a simple formula : $(n + 2) * 2^{n-1}$.

The number of compositions with a given length is given by $C(n - 1, k)$, where k is the $length - 1$.

Here is the table for n=5:

Note that the formula quoted in this section are only applicable to the full version of compositions, and only hold to $n = 9$ in this paper.

§4. The function $NND(y, X)$

In order to perform our calculations, we need to first determine an X which is our upper bound, and we do not include X. We then define a function $Number_Of_Numbers_With_Digit_Sum_d(X)$, or $NND(y, X)$ for short.

We count the number of digits of X, and this is the maximum length that we need to consider.

For simplicities sake, we will only consider X to be a power of 10.

If $X = 10^z$, then $NND(1, X) = z$, and as the maximum digit sum is $9 * (z - 1), NND(9 * (z - 1)) = 1$.

Next we consider $NND(2, X)$ through to $NND(9, X)$, as these obey the formulas given in the previous section.

$NND(2, X)$

$NND(2, X)$ asks for the number of integers with a digit sum of 2 on the range 1 to 10^z.

This requires either an opening digit of 2, and the rest zeroes, or an opening digit of 1, some zeroes, another 1, and some more zeroes, possibly zero zeroes in each case.

There are z examples of the first case, for example if $z = 3$ we have $2, 20$ and 200.

The second case we consider the stretched partitions $11, 101, 1001, 10001$, etc..., as stems for the numbers we are interested in.

If $z = 3$, then the number of each stem present is $2[11, 110]$ and $1[101]$.

If $z = 4$, then we have $3[11, 110, 1100]$ plus $2[101, 1010]$ plus $1[1001]$.

Therefore $NND(2, X) = z(z + 1)/2$.

$NND(3, X)$

We have 4 cases:

I. 3

II. 21

III. 12

IV. 111

Case I : Easy - contributes z

Case II and III : contributes $(z - 1)z/2$ once each - total $z(z - 1)$

Case IV : We have two gaps (a stem must end with a 1), and this creates k possibilities for each stretched composition into $k + 2$ digits;

Stems: $111, [1011, 1101], [10011, 10101, 11001], [100011, 100101, 101001, 110001], \ldots$

To gain some idea of how this pans out, we create a table:

These are the tetrahedral numbers [5] given by $(z - 2)(z - 1)(z)/6$

So we have

$$z + z^2 - z + z^3/6 - z^2/2 + z/3 = z^3/6 + z^2/2 + z/3$$
$$= (z^3 + 3z^2 + 2z)/6 = (z)(z + 1)(z + 2)/6$$

This predicts $2 * 3 * \frac{4}{6} = 4$ entries less than 100, and these are $3, 12, 21, 30$. The 10 entries less than 1000 include these 4 and also $102, 111, 120, 201, 210$ and 300.

$NND(4, X)$ **to** $NND(9, X)$

From the previous results, we might reasonably expect that for y between 0 and 9 we have:

$$NND(y,x) = \frac{\prod_{j=0}^{y-1} z+j}{y!}$$

Before we attempt to prove this, we should test our hypothesis:

First construct a table of z values to say $z = 6$ and the predicted value of $NND(y, X)$:

Define a function $PNND(y, z)$ as in the equation above:

$PNND(y, z) = 1/y! * prod(j = 0, y - 1, z + j)$

And run through y and z to create the table above.

To physically determine the table use:

$DNND(y, z) = local(c); c = 0; for(i = 1, 10^z - 1, if(sd(i) == y, c++)); c$

which produces the desired output (8 minutes at $3Ghz$).

§5.Proof

We can see how this might be true due to the nature of the table. Each row (or column) is the partial sums of the previous row (or column).

This leads us to consider that the answer for a digit sum is constructible from the previous digit sum if we read the candidates in a different way.

For example, the $y = 1$ row gives $[1], [1, 10], [1, 10, 100], etc\ldots$

Therefore for $y = 2$, consider the grouping:

[1],

[[1], [1,10]],

[[1], [1,10], [1,10,100], etc...

Can we interpret this as $[2], [2, 11, 20], [2, 11, 20, 101, 110, 200], etc\ldots$?

And can we find a mapping between the values at $y = 3$, namely [3], $[3, 12, 21, 30]$, $[3, 12, 21, 30, 102, 111, 120, 201, 210, 300]$ and the groups formed from $y = 2$ under our hypothesis 2?

And can we find a method of extending this through to $n = 9$?

A logical map involves adding 1 to each element, and then extending elements in group with zeroes depending on the relative position of the group inside the main group.

So looking at the first case, [1] becomes [2].

[1] becomes [20] as it is a 1 digit value required to become a 2 digit value. $[1, 10]$ becomes $[2, 11]$

Next, [1] becomes [200], $[1, 10]$ becomes $[20, 110]$ and $[1, 10, 100]$ becomes $[2, 11, 101]$.

This works for 3 as well, but we need to justify the argument up to the map between $y = 8$ and $y = 9$.

And we can do this by first observing that the number of trailing zeroes makes each group unique, and within each group each new element is obviously unique.

$NND(10+, X)$

With y greater than 9, the pattern stops, as for example with $NND(10, X)$ we cannot have 10 as a composition.

The formula predicts $1, 11, 66, 286, 1001, 3003$ and $DNND(10, z)$ states that the actual number of numbers with a digit sum 10 is $0, 9, 63, 282, 996, 2997$

So $NND(10, X) = predicted - z$, due to the missing composition.

With $NND(11, X)$, we have neither 11 or $10+1$ or $1+10$ as compositions.

Predicted is $1, 12, 78, 364, 1365, 4368$ Actual is $0, 8, 69, 348, 1340, 4332$

Which leads us to believe $NND(11, X) = predicted - z^2$

With $NND(12, X)$;

Predicted is $1, 13, 91, 455, 1820, 6188$ Actual is $0, 7, 73, 415, 1745, 6062$

The difference here is the pentagonal pyramidal numbers [6], so

$NND(12, X) = predicted - z^2(z + 1)/2$

There seems to be a pattern, so let's continue for a while:

NND(13,X) Predicted : $1, 14, 105, 560, 2380, 8568$ Actual : $0, 6, 75, 480, 2205, 8232$

$NND(13, X) = predicted - z^2(z + 1)(z + 2)/6$ [7]

$NND(14, X)$ Predicted : $1, 15, 120, 680, 3060, 11628$

Actual :$0, 5, 75, 540, 2710, 10872$

$NND(14, X) = predicted - (z + 1)^2(z + 2)(z + 3)(z + 4)/24$ [8]

The pattern is reasonably obvious, but unstable - we cannot say when, if ever, a jump to $(n + 2)^2$ as a start occurs.

If we put the data we have into a table, and incorporate the previous table for $y = 0$ to 9, we might spot a pattern:

And we do have a $pattern - each$ column is symmetric!, i.e. after $y = 9$, column 2 is itself reversed, after $y = 13$, column 3 is itself reversed, and (we guess) so on. The switch happens at $y = 9z/2$, which is the maximum sum of digits $(i.e. 99 \ldots 99)$ divided by 2.

The symmetry can be seen by considering that if a number n has a digit sum d, then $(10^z - 1) - n$ has a digit sum $9z - d$.

And we have another $pattern - after$ $y = 9$, a column continues to be the partial sums of the previous column, however the $y - 10'th$ entry is subtracted.

For example, consider the $z = 4$ column. $(9, 4) = 220$, and $(10, 4) = (9, 4) + (10, 3) - (0, 3) = 220 + 63 - 1 = 282$. $(11, 4) = (10, 4) + (11, 3) - (1, 3) = 282 + 69 - 3 = 348$.

In general $(y, z) = (y-1, z) + (y, z-1) - (y-10, z-1)$, where $(y, z) = 0$ if y is less than 0.

This becomes obvious if we see that the stretched compositions not present in y greater than 9 are equal in number to the number of stretched compositions of $y - 10$.

Summary

We prove that $NND(y, X)$ for $X = 10^z$ and y less than 10 is given by product$(j = 0, y - 1, z + j)/y!$.

We show that $NND(y, X)$ with y greater than 9 behaves reasonably predictably.

We prove that $NND(y, X) = NND(9z - y, X)$

We prove that $(y, z) = (y - 1, z) + (y, z - 1) - (y - 10, z - 1)$

Open problems

1. Find a formula for $NND(y, X)$ valid for all y with $X = 10^z$
2. Find a formula for $NND(y, X)$ valid for y less than 10 with a general X
3. Find a formula for $NND(y, X)$ valid for all y with a general X

References

[1] SOME NOTIONS AND QUESTIONS IN NUMBER THEORY Sequence, 16 - Smarandache Digital Sums, http://www.gallup.unm.edu/ smarandache/SNAQINT.txt

[2] Pari/GP http://pari.math.u-bordeaux.fr/

[3] A007953 http://www.research.att.com/projects/OEIS?Anum=A007953

[4] A001792, OEIS

[5] A000292, OEIS

[6] A002411, OEIS

[7] A002417, OEIS

[8] A027800, OEIS

[9] A056003, OEIS

[10] A027800, OEIS