Finding the k-th rational number in Cantor enumeration: a closed formula

Massimo Zanetti

Abstract

Here, we consider the Cantor enumeration of rational numbers and we provide the analytical relationship between the indexing variable k and the numerator/denominator of the k-th rational number. We see that this relationship can be given in terms of simple operations: one square root and a few multiplications and sums. The theoretical argument makes use of triangular numbers and some of their simple properties.

Problem formulation and solution

Let us consider the Cantor enumeration of rational numbers. For each $k \in \mathbb{N}^*$ the k-th rational number $r(k)$ is listed as in

k	1	2	3	4	5	6	7	8	9	10	11	12	...
$r(k)$	$\frac{1}{1}$	$\frac{2}{1}$	$\frac{1}{2}$	$\frac{3}{1}$	$\frac{2}{2}$	$\frac{1}{3}$	$\frac{4}{1}$	$\frac{3}{2}$	$\frac{2}{3}$	$\frac{1}{4}$	$\frac{5}{1}$	$\frac{4}{2}$	...

Our problem can be stated as follows.

Problem. *Given $k \in \mathbb{N}^*$, find both the numerator and the denominator of the k-th rational number $r(k)$ as functions of k.*

To provide a solution of the problem which is computationally convenenient, we start by firstly observing that the Cantor enumeration can be better visualized by representing numerators and denominators as rows and columns indeces of a matrix whose entries are the indeces of the sequence:

	1	2	3	4	5	6	7	8	9	10	11	12	...
1	1	2	4	7	11	16	...						
2	3	5	8	12	17	...							
3	6	9	13	18	...								
4	10	14	19	...									
5	15	20	...										
6	21	...											
7	...												
$\vdots$	...												

To get the k-th rational number in an inexpensive way, we exploit triangular numbers. The reason for this comes from an equivalent formulation of our problem, that we show here. Looking at sequence of numerators N and denominators

D of the Cantor enumeration

$$\begin{array}{c|cccccccccccccccccc} k & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & \dots \\ N(k) & 1 & 2 & 1 & 3 & 2 & 1 & 4 & 3 & 2 & 1 & 5 & 4 & 3 & 2 & 1 & 6 & 5 & \dots \\ D(k) & 1 & 1 & 2 & 1 & 2 & 3 & 1 & 2 & 3 & 4 & 1 & 2 & 3 & 4 & 5 & 1 & 2 & \dots \end{array}$$

we can see they follow a triangular pattern that we can list in three matrices

$$I = \begin{bmatrix} 1 & & & & \\ 2 & 3 & & & \\ 4 & 5 & 6 & & \\ 7 & 8 & 9 & 10 & \\ 11 & 12 & 13 & 14 & 15 \\ \dots & & & & \end{bmatrix} N = \begin{bmatrix} 1 & & & & \\ 2 & 1 & & & \\ 3 & 2 & 1 & & \\ 4 & 3 & 2 & 1 & \\ 5 & 4 & 3 & 2 & 1 \\ \dots & & & & \end{bmatrix} D = \begin{bmatrix} 1 & & & & \\ 1 & 2 & & & \\ 1 & 2 & 3 & & \\ 1 & 2 & 3 & 4 & \\ 1 & 2 & 3 & 4 & 5 \\ \dots & & & & \end{bmatrix}.$$

Given this representation, one can easily see that given $k \in \mathbb{N}^*$, the knowledge of the indeces $x, y \in \mathbb{N}^*$ such that $I(x, y) = k$ allows us to uniquely represent the rational number $r(k)$ as

$$r(k) = \frac{N(x, y)}{D(x, y)}.$$

Moreover, we observe that the rows of N can be obtained by flipping the rows of D and the row index is retained by the first entry in each row of N. Therefore, we have that

$$r(k) = \frac{x - y + 1}{y}. \tag{1}$$

These facts allows us to formulate the following.

Problem (Equivalent formulation)**.** *Given $k \in \mathbb{N}^*$, find the indeces $x, y \in \mathbb{N}^*$ such that $I(x, y) = k$.*

In order to see why triangular numbers are useful to the solution of our problem, let us first recall their definition.

Definition (Triangular numbers)**.** *A natural number $n \in \mathbb{N}^*$ is triangular if there exists $g \in \mathbb{N}^*$ such that*

$$n = \sum_{\ell=1}^{g} \ell.$$

If this is the case, we say that g is the tri-generator of n, and, given the well-known formula for the summation of the first g natural numbers, we have that

$$n = \sum_{\ell=1}^{g} \ell = \frac{g(g+1)}{2}. \tag{2}$$

To emphasize that g is the tri-generator of n, we also write g^n.

Given a triangular number n, we can find g^n by analyzing the two solutions of equation

$$g^2 + g - 2n = 0, \tag{3}$$

which is easily obtained from (2). The two solutions are

$$g_1 = \frac{-1-\sqrt{1+8n}}{2}, \qquad g_2 = \frac{-1+\sqrt{1+8n}}{2}$$

and, being $g_1 < 0$ and $g_2 > 0$ for any n, we have that $g^n = g_2$ (this can also be proved by induction).

Case 1. We notice that, if k is triangular number it lies on the diagonal of the I matrix and its tri-generator $g^k = (-1+\sqrt{1+8k})/2$ is both its row and column index, i.e., $x = y = g^k$ and $I(x,y) = k$.

Case 2. Let us now assume that $k > 1$ is not a triangular number. Then, we consider the two subsequent triangular numbers k_1, k_2 such that $k_1 < k < k_2$. By the definition of tri-generator and by strict monotonicity of the parabola in its rigth branch, we easily have that

$$g^{k_1} < g^k < g^{k_2}$$

with $g^{k_1}, g^{k_2} \in \mathbb{N}^*$ and $g^{k_2} = g^{k_1} + 1$. This implies that the raw index x of k in matrix I is $x = g^{k_2} = \lceil g^k \rceil$, where $\lceil . \rceil$ is the ceiling function. Given x, finding the column index y such that $I(x,y) = k$ is trivial. In fact, because of the lower diagonal shape of matrix I, the row x has exactly x terms spanning columns from 1 to x and ranging from values $k_1 + 1$ to k_2. If follows that we can obtain y (the column place of k in this raw) by subtracting the distance between k_2 and k to x, i.e.,

$$y = x - (k_2 - k).$$

By writing k_2 in terms of its tri-generator g^{k_2} and by recalling that $x = g^{k_2}$, we get:

$$y = x - \frac{g^{k_2}(g^{k_2}+1)}{2} + k = \frac{-x^2+x}{2} + k.$$

Finally, we note that this expression for y is valid also in the case k is a triangular number. Indeed, as previously seen $x = g^k$ is tri-generator of k, thus $k = x(x+1)/2$. Therefore, we get

$$y = \frac{-x^2+x}{2} + k = \frac{-x^2+x}{2} + \frac{x(x+1)}{2} = x.$$

Plugging the two expressions for x, y into (1), we get the analytical expression of the k-th rational number in Cantor enumeration in terms of k.

Lemma (Cantor enumeration of rational numbers)**.** *Given $k \in \mathbb{N}^*$, the k-th rational number in the Cantor enumeration is*

$$r(k) = \frac{\frac{x(x+1)}{2} - (k-1)}{\frac{x(-x+1)}{2} + k}.$$

where $x = \lceil(-1+\sqrt{1+8k})\rceil$.

Numerical implementation

Here is the function to compute the k-th rational number in the Cantor enumeration using Matlab, with some examples of the results.

```
function kthRationalCantor(k)
    x = ceil((-1+sqrt(1+8*k))/2);
    y = ((-x^2+x)/2)+k;
    fprintf('The %d-th rational number in Cantor enumeration is: ...
        %d / %d \n',k,x-y+1,y);
end
```

```
>> kthRationalCantor(6)
The 6-th rational number in Cantor enumeration is: 1 / 3
>> kthRationalCantor(385)
The 385-th rational number in Cantor enumeration is: 22 / 7
>> kthRationalCantor(108924)
The 108924-th rational number in Cantor enumeration is: 355 / 113
>> kthRationalCantor(2085704)
The 2085704-th rational number in Cantor enumeration is: 200 / 1843
```