A Simple Method for Calculations of the Number of Inversions in Permutation

Dominik Strzałka[1*]

[1]*Rzeszów University of Technology, Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland.*

Author's contribution

The sole author designed, analyzed and interpreted and prepared the manuscript.

Short Research Article

ABSTRACT

The aim of this paper is to show the recurrence method for obtaining the number of inversions $I_n(k)$ in input sets with different sizes n, when the information about $I_{n-1}(k)$ is given. The proposed method is based on a simple observation that the use of recursive approach gives an elegant way for obtaining those numbers in contradiction to the so far existing approach based on binomial coefficients and pentagonal numbers. The complexity of this method is $O(n^3)$. The results of this proposal can be used for interesting exercises in education of maths and also for problem of inversions description in sorting algorithms.

*Corresponding author: E-mail: strzalka@prz.edu.pl;

1 INTRODUCTION

In combinatorics, a permutation is usually understood as a sequence containing each element from a finite set once and only once. Due to the fact that permutations can also represent the unsorted set of elements, in many cases their properties can be used in the analysis that considers the problem of sorting algorithms performance in computer science.

Let $a_1 a_2 \ldots a_n$ denote the permutation of n-element set $\{1, 2, \ldots, n\}$. Each pair (a_i, a_j) that fulfills the condition $a_i > a_j$ for $i < j$ is called an inversion (see [1]). For permutation: $\{6 \ 7 \ 1 \ 2\}$ the following set of inversions can be given: $\{(6,1), (6,2), (7,1), (7,2)\}$. In computer science the existence of at least one inversion in input set denotes that this set is unsorted and that is why this term is often used in the analysis of sorting algorithms. For example, let's consider the sequence $\{1, 2, 7, 6\}$ where the only one inversion $(7,6)$ exist. Its existence means that the sequence is unsorted. The term *inversion* was first introduced by Gabriel Cramer in 1750 (see [2]), who took into account the problem of finding the solution for linear equation by matrix determinants.

The table of inversions $b_1 b_2 \ldots b_n$ for permutation $a_1 a_2 \ldots a_n$ can be obtained by the assumption that b_j is the number of such elements from sequence that are on left side of j, which are greater than j, thus b_j is the number of inversions, whose second component is j (see [3]). For example, consider the following permutation:

$$6 \quad 7 \quad 1 \quad 2 \quad 4 \quad 9 \quad 5 \quad 8 \quad 3 \qquad (1.1)$$

which has the inversion vector

b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
2	2	6	2	3	0	0	1	0

With regard to the Marshall Hall theorem ([4]) it is worth noting that the inversion vector explicitly defines only one corresponding permutation. For sequence (1.1) it can be achieved in the following manner:

Number 9 should be the first written number. Then number 8, because from inversion vector it can be seen that it has only one inversion, thus it should be written on right side of 9. Next, number

7 – in inversion vector $b_7 = 0$ thus it should be written on the left side of 9. Similarly 6, because $b_6 = 0$, thus 6 should be on left side of 7. The sequence 6 7 9 8 has been obtained so far. For 5, $b_5 = 3$ thus 5 should be written between 9 and 8, because it guarantees the existence of three inversions, but 4 should be written between 7 and 9 because this guarantees the existence of 2 inversions ($b_4 = 2$), and so on. As a result the sequence (1.1) is obtained.

It can be very easily verified that when in one sequence two neighboring elements are changed the number of inversions raises or falls by 1.

2 NUMBER OF INVERSIONS

A very interesting issue is the question of how many permutations with k inversions are present in n-element set. If $n = 1$ then there are no inversions, thus $k = 0$. For $n = 2$ there are only two possible cases – $k = 0$ (all elements are in the right order) or $k = 1$ (there is one inversion $a_2 a_1$).

For $n = 3$ the following are possible: one case $(a_1 a_2 a_3)$ with $k = 0$, two cases $(a_1 a_3 a_2, a_2 a_1 a_3)$ with $k = 1$ inversion and inversion vectors $(b_1 b_2 b_3 = 0, 1, 0; b_1 b_2 b_3 = 1, 0, 0)$, next two cases $(a_2 a_3 a_1, a_3 a_1 a_2)$ with $k = 2$ inversions and inversion vectors $(b_1 b_2 b_3 = 2, 0, 0; b_1 b_2 b_3 = 1, 1, 0)$ and also one case $(a_3 a_2 a_1)$ with inversion vector $(b_1 b_2 b_3 = 2, 1, 0)$. The calculation of next cases requires consideration of $n!$ permutations. If the number of k inversions in n-element set will be denoted by $I_n(k)$ consequently, taking into account the vector $b_1 b_2 \ldots b_n$ it can be observed that $I_n(0) = 1, I_n(1) = n - 1$ and

$$I_n \left(\binom{n}{k} - k \right) = I_n(k). \qquad (2.1)$$

In Table 1 all numbers under the emphasized (bold) positions fulfill

$$I_n(k) = I_n(k-1) + I_{n-1}(k), \qquad \text{for } k < n \qquad (2.2)$$

It can be shown that (see [5]):

$$
\begin{aligned}
I_n(2) &= \binom{n}{2} - 1 & \text{for } n \geq 2 \\
I_n(3) &= \binom{n+1}{3} - \binom{n}{1} & \text{for } n \geq 3 \\
I_n(4) &= \binom{n+2}{4} - \binom{n+1}{2} & \text{for } n \geq 4 \\
I_n(5) &= \binom{n+4}{5} - \binom{n+2}{3} + 1 & \text{for } n \geq 5
\end{aligned}
\qquad (2.3)
$$

and generally the equation for calculating $I_n(k)$ can be explained as

$$I_n(k) = \binom{n+k-2}{k} - \binom{n+k-3}{k-2} + \binom{n+k-6}{k-5} + \binom{n+k-8}{k-7} \quad - \quad \ldots$$

$$+(-1)^j \left(\binom{n+k-u_j-1}{k-u_j} + \binom{n+k-u_j-j-1}{k-u_j-j} \right) \quad + \quad \ldots \quad n \geq k \quad (2.4)$$

where $u_j = (3j^2 - j)/2$ and is called the pentagonal number (see Euler's Pentagonal Number Theorem in [6]).

Table 1 presents the probability distribution of number of inversions in n-element permutation. In order to illustrate this it is enough to divide each number in a given row of Table 1 by $n!$ What can be noted from this is that the average number of inversions equals (see [5]):

$$\mu(n) = 0 + \frac{1}{2} + \frac{3}{2} + \ldots + \frac{n-1}{2} = \frac{n(n-1)}{4} = \frac{n^2-n}{4}, \tag{2.5}$$

while the variance can be computed from

$$\sigma^2(n) = 0 + \frac{1}{4} + \frac{5}{6} + \ldots + \frac{n^2-1}{12} = \frac{n(2n+5)(n-2)}{72}. \tag{2.6}$$

It means that the standard deviation is proportional to

$$\sigma(n) \propto \frac{1}{6} n^{\frac{3}{2}}. \tag{2.7}$$

The numbers $I_n(k)$ can be used for example in computational complexity average case analysis in insertion sort algorithm because the number of inversions in permutation $(a_1 a_2 \ldots a_n)$ is used for calculations of the total average time (expressed by the number of dominant operations) needed to sort n-element input set (examples are given in [7]).

Table 1: Number $I_n(k)$ inversions in n-th element permutation

n	$I_n(0)$	$I_n(1)$	$I_n(2)$	$I_n(3)$	$I_n(4)$	$I_n(5)$	$I_n(6)$	$I_n(7)$	$I_n(8)$	$I_n(9)$	$I_n(10)$
1	**1**	0	0	0	0	0	0	0	0	0	0
2	1	**1**	0	0	0	0	0	0	0	0	0
3	1	2	**2**	1	0	0	0	0	0	0	0
4	1	3	5	**6**	5	3	1	0	0	0	0
5	1	4	9	15	**20**	22	20	15	9	4	1
6	1	5	14	29	49	**71**	90	101	101	90	71

3 SIMPLE RECURRENCE METHOD

On the basis of the data gathered in Table 1 it is clear to notice that the frequency of occurrence k inversions in n element set undergoes very simple relations. It can be defined in the following manner. Let $n = 3$, thus there are 6 permutations and the number of k inversions can be seen in Table 1. If there is a need to calculate the number of possible inversions for $n = 4$ (there are 24 permutations) it can be done considering the information for $n - 1 = 3$. The following operation should be carried out: sum $n = 4$ times the number of inversions for $n-1 = 3$ however for each partial sum move the numbers one column to the right. The example is shown in Table 2 and in Fig. 1.

Table 2: Example of the method

$n=3$	1	2	2	1							
		1	2	2	1						
			1	2	2	1					
				1	2	2	1				
$n=4$ Σ	1	3	5	6	5	3	1				
		1	3	5	6	5	3	1			
			1	3	5	6	5	3	1		
				1	3	5	6	5	3	1	
					1	3	5	6	5	3	1
$n=5$ Σ	1	4	9	15	20	22	20	15	9	4	1

Figure 1: Example of the method

The method presented above is based on recurrence approach. Calculations for the given n can be done when the information about $n-1$ is given. However, in each recurrence call of method there is a need of n-times adding of $k(n)$ possible inversions where:

$$k(n) = k(n-1) + n - 1 \qquad \text{where} \qquad k(1) = 1. \tag{3.1}$$

The solution of equation (3.1) is:

$$k(n) = 1 + \sum_{i=1}^{n}(i-1) = 1 + \frac{n(n-1)}{2} \tag{3.2}$$

thus $k(n)$ is $O(n^2)$.

Since the presented method requires n times recall of $k(n)$, its time complexity is $O(n^3)$.

4 CONCLUSIONS

The key aspect of this work was to outline the simple recurrence method for obtaining the $I_n(k)$ numbers. Its complexity is polynomial and method doesn't require any additional calculations unlike in the case of the method based on binomial coefficients where the pentagonal numbers are used. This method can be very useful in the analysis of sorting algorithms properties and a very easy way to measure the extent to which the permutation is *out of order*. Moreover, thanks to presented explanations it is possible to use this method in education of maths.

COMPETING INTERESTS

The author declares that no competing interests exist.

References

[1] Margolius BH. Permutations with inversions. Journal of Integer Sequences. 2001;4 art. 01.2.4.

[2] Muir T. The theory of determinants in the historical order of development. Macmillan, London; 1906.

[3] Skiena S. Inversions and inversion vectors. §1.3 in Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica. Reading, MA: Addison-Wesley. 1990;27-31.

[4] Hall M. Jr. Combinatorial theory. Blaisdell Pub. Co. Waltham; 1967.

[5] Knuth D. The art of computer programming. Addison-Wesley; 1976.

[6] Andrews GE. Euler's pentagonal number theorem. Mathematics Magazine. 1983;56(5):279-284.

[7] Gassner BJ. Sorting by replacement selecting. Communications of the ACM. 1967;10(2):89-93.