Combinatorial characterization of inrankings as weak orders induced by intervals

S V Muravyov and E Y Emelyanova
National Research Tomsk Polytechnic University
Pr. Lenina, 30, Tomsk, 634050, Russian Federation

E-mail: muravyov@tpu.ru

Abstract. The problem of reliable processing of heteroscedastic interval data occupies an important niche among urgent topics of measurement science. The paper is devoted to a combinatorial characterization of so called ‘inrankings’ which are weak orders induced by input intervals of the interval fusion with preference aggregation (IF&PA) procedure. The procedure transforms the given \( m \) initial real line intervals into inrankings, which are a specific case of weak order relations (or rankings) over a set of \( n \) discrete values belonging to these intervals. The new notation of inranking appears as a result of restrictions imposed on the ordinary rankings by interval character of the initial data. In the paper, the inranking spaces properties are investigated from the combinatorial theory point of view. It is shown that the inranking space is a subset of the set of all weak orders with a single symbol of strict order. The cardinality of inranking space is defined by the triangle number for the given number \( n \) of the discrete elements. Cardinalities of other adjacent spaces are considered.

1. Introduction
The role of interval data integration and interpretation both in theory and practice of measurement data fusion can hardly be overestimated [1]. Suffice it to recall an importance of the confidence interval parameters estimation in a traditional statistical framework with various applications both in physical and social sciences [2]. In this paper we deal with the problem of heteroscedastic interval data processing. Heteroscedasticity means observations heterogeneity manifested as changeable variance of a regression model random error [3]. The random errors heteroscedasticity results in inefficiency of estimates obtained using the least square method. In turn, this can lead to inadequate statistical conclusions on estimated quantity values and their quality.

Generally, the data fusion refers to a joint processing of data on some object obtained from multiple sources aiming to acquire fuller, more objective and accurate knowledge of a characteristic under investigation than knowledge derived from a single source [1]. The interval data fusion (figure 1) can be deemed as a procedure intended to find an interval to be consistent with maximal number of the initial intervals and with maximal likelihood including a
representing them value \( x^* \). Existing methods of the heteroscedastic interval data fusion, such as combination of weighted mean and \( \chi^2 \)-test [3], different versions of approval voting [4], etc. do not provide simultaneously a uniqueness of the fusion outcome and its robustness in terms of an independence on particular law of the interval data probability distribution. In [5], the paper authors have been proposed the interval fusion with preference aggregation (IF&PA) and have demonstrated the method outcomes’ guaranteed improved robustness and accuracy by Monte Carlo experimentation. The IF&PA procedure has been successfully applied in inter-laboratory comparison data processing, enhancement of multisensor readings accuracy and providing energy-accuracy trade-off in presence of outliers in wireless sensor network [6, 7, 8].

The IF&PA transforms the given initial real line intervals into inrankings, which are a specific case of weak order relations (or rankings) over a set of discrete values belonging to these intervals. The inrankings appears as a result of restrictions imposed on the ordinary rankings by interval character of the initial data. Then, for the preference profile consisting of the inrankings, the IF&PA determines a single consensus weak order relations (or rankings) over a set of discrete values belonging to these intervals. The inrankings specification a binary relation of weak order on the set \( A \) [10]. It is easy to show that \( \rho = \lambda^{-1} \) and \( \tau = \lambda \cap \lambda^{-1} = \rho \cup \rho^{-1} \). Set \( \Lambda(n, m) = \{ \lambda_1, \lambda_2, \ldots, \lambda_m \} \) of \( m \) rankings of \( n \) elements \( a_i, i = 1, \ldots, n \), is called a preference profile for \( m \) and \( n \) given.

\[
\{I_k\}, k = 1, \ldots, m
\]

\[
\begin{array}{c}
I_1 \quad \cdots \quad I_k \quad \cdots \quad I_n \\
\text{Discrete values} \\
\{a_1, a_2, \ldots, a_{n-1}, a_n\}
\end{array}
\]

\[
\begin{array}{c}
\text{Range of actual values formation} \\
[a_1; a_n] \\
\text{Preference profile \( \Lambda \) formation} \\
\Lambda(n, m)
\end{array}
\]

\[
\begin{array}{c}
\text{Range of actual values partition} \\
\lambda_1: a_1 \sim a_2 \sim a_3 \sim a_4 \sim \cdots \sim a_5 \\
\cdots \\
\lambda_m: a_4 \sim a_5 \sim a_1 \sim a_2 \sim a_3 \\
\end{array}
\]

**Figure 2.** Stages of conversion of a set of intervals to a preference profile.

Consider a collection of \( m \) closed intervals \( \{I_k\}_{k=1}^m \) in the real line, where each interval has a *middle point* \( x_k \) a *lower bound* \( x_k - \varepsilon_k \) and *upper bound* \( x_k + \varepsilon_k \), so that \( I_k = [x_k - \varepsilon_k, x_k + \varepsilon_k]; x_k, \varepsilon_k \in \mathbb{Y} \). Introduce a range of actual values (RAV) \( A = \{a_1, a_2, \ldots, a_n\} \), over which the strict order (inherited from the real line) \( a_1 < a_2 < \cdots < a_n \) exists. The discrete set \( A \) can be produced starting from the given continuous intervals \( \{I_k\}_{k=1}^m \) in three steps shown in figure 2. First, RAV’s lower and upper bounds are chosen, i.e. \( a_1 = \min \{x_k - \varepsilon_k \mid k = 1, \ldots, m\} \) and \( a_n = \max \{x_k + \varepsilon_k \mid k = 1, \ldots, m\} \) correspondingly. Then, to generate elements \( a_2, a_3, \ldots, a_{n-1} \), the obtained interval \( [a_1; a_n] \) is partitioned into \( n - 1 \) equal subintervals of length \( h = (a_n - a_1)/(n - 1) \), and the \( i \)-th RAV’s element becomes \( a_i = a_{i-1} + h, i = 2, \ldots, n \). Finally, the set \( A \)
= \{a_1 < a_2 < \ldots < a_n\} of strictly ordered discrete values \(a_i, i = 1, \ldots, n\), is used to shape the preference profile \(\Lambda(n, m)\) representing the initial intervals \([I_k^m]\) \cite{5}. The conditions for formation of the profile are discussed at the beginning of next section.

3. Inrankings and inranking spaces

For any interval \(I_k, k = 1, \ldots, m\), we have \(A = A_k \cup \tilde{A}_k\), \(A_k \cap \tilde{A}_k = \emptyset\), where the subset \(A_k\) includes all the elements of \(A\) belonging to the interval \(I_k\), and its complement \(\tilde{A}_k\) includes all the rest elements of \(A\). The ranking induced by interval \(I_k\) will be called the inranking \(\lambda_k\), if it satisfies the following four conditions for \(i, j = 1, \ldots, n\):

\[a_i \in A_k \land a_j \notin A_k \Rightarrow a_i > a_j;\]  
\[a_i, a_j \in A_k \lor a_i \notin A_k \Rightarrow a_i > a_j;\]  
\[a_i \notin A_k \land a_j \in A_k \Rightarrow a_i < a_j;\]  
\[a_i, a_j \in A_k \land \text{are neighbors} \Rightarrow j = i + 1.\]  

Notice that the \(k\)-th ranking consists of two equivalence classes composed of elements of the sets \(A_k\) and \(\tilde{A}_k\). Elements of the class \(A_k\) are strictly preferred over elements of the class \(\tilde{A}_k\), i.e. always \(\lambda_k = A_k \succ \tilde{A}_k\).

Hence, the ranking contains a single symbol of strict order \(\succ\) and \(n - 2\) symbols of tolerance \(\sim\). Example 1.

One of possible ranking for \(n = 5\) could be \(\lambda_4 = \{a_2 \sim a_3 \succ \sim a_4 \succ a_5\}, \text{ where } A_k = \{a_2 \sim a_3 \sim a_4\}, \text{ and } \tilde{A}_k = \{a_1 \sim a_3 \succ a_5\}.\)

For notational convenience, ranking elements will be denoted by their indexes, and subsets \(A_k\) and \(\tilde{A}_k\) will be shown by parenthesizing appropriate sets of indexes. Then the symbols \(\succ\) and \(\sim\) are omitted.

Example 2. The inranking from example 1, \(\{a_2 \sim a_3 \succ a_1 \succ a_5\}\), can be concisely represented as \((234)(15)\). □

The condition (4) is necessary as series of elements \(\{a_i\}\) of the set \(A\) is strongly monotonic, i.e., \(a_i < a_{i+1}\) for all \(i \in \mathbb{N}\). The class \(A_k \subseteq A\) can include merely sequential subseries of elements of \(A\) without breaks, that is, indexes of these elements must constitute a segment of the natural series. It means that the difference of indexes for any two neighbor elements \(a_i\) and \(a_j\) in \(A_k\) cannot be larger than 1, i.e., \(j = i + 1\).

A concrete preference profile always consists of inrankings taken from the space of all possible ones. Hence, now we try to answer the question how many different inrankings exist for given \(n\).

Let us denote the space of weak orders via \(\Omega\) and the space of inrankings via \(\tilde{\Omega}\). We need also to consider the space \(\Omega_i\) elements of which satisfy the conditions (1)–(3) but violate the condition (4). In fact, the members of \(\Omega_i\) are just those having a single symbol of strict order \(\succ\). Clear that

\[\Omega_2 \subseteq \Omega_1 \subseteq \Omega_0,\]  

for which the following expressions are valid:

\[\Omega_1 = \Omega_2 \cup (\tilde{\Omega}_2 \cap \Omega_1), \quad \Omega_1 \cap \tilde{\Omega}_1 = \emptyset, \quad \Omega_2 \cap \tilde{\Omega}_2 = \emptyset,\]

\[|\Omega_1| = |\Omega_2| + |\tilde{\Omega}_2 \cap \Omega_1|.\]  

The space \(\Omega_0\) plays the part of universal set for the sets \(\Omega_1\) and \(\Omega_2\). The space \(\Omega_1\) consists of the space \(\Omega_2\) of inrankings and the space of forbidden rankings \(\tilde{\Omega}_1 = \tilde{\Omega}_2 \cap \Omega_1\).

Cardinalities of the space \(\Omega_0\) and its subspaces \(|\Omega_1|\) and \(|\Omega_2|\) are connected to Stirling numbers of the second kind \(S_{n,q}\) that defines number of unordered partitions of \(n\)-element set into \(q\) non-empty subsets \cite{11}:

\[|\Omega_0| = \sum_{q=0}^{n} q! S_{n,q},\]  

where

\[\text{Cardinalities of the space } \Omega_0 \text{ and its subspaces } |\Omega_1| \text{ and } |\Omega_2| \text{ are connected to Stirling numbers of the second kind } S_{n,q} \text{ that defines number of unordered partitions of } n \text{-element set into } q \text{-non-empty subsets } \text{[11]}:\]
Composition of the spaces $\Omega_1$, $\Omega_2$ and $\Omega_2 \cap \Omega_1$ for $n = 1, \ldots, 5$, at corresponding cardinalities of $\Omega_0$ and $A_k$, is shown in table 1. It is clear from table 1 that cardinality of the in-ranking space $|\Omega_2|$ in dependence on $n$ follows the series 1, 3, 6, 10, 15, 21, 28, 36, 45, 55, ..., elements of which are called triangle numbers $T_n$ [12]. A triangle number and, hence, the cardinality $|\Omega_2|$ are defined by formula:

$$\frac{n(n + 1)}{2}.$$ 

| $\Omega_0$ | $|A_k|$ | $\Omega_1$ | $|\Omega_1|$ | $\Omega_2$ | $|\Omega_2|$ | $\Omega_f$ | $|\Omega_f|$ | $n - |A_k| + 1$ |
|-----------|--------|-----------|-------------|-----------|-------------|-----------|-------------|------------------|
| $n = 1$   |        | 1         | 1           | 1          | 1           | $\emptyset$ | 0            | 1                |
| $n = 2$   | 3      | 12, 21    | 3           | 12, 21     | $\emptyset$ | 0          | 1            | 2                |
| $n = 3$   | 13     | 1(23), 2(13), 3(12) | 7 | 1(23), 2(13), 3(12) | (123) | 6 | $\emptyset$ | 2 | 3 |
| $n = 4$   | 75     | 1(234), 2(134), 3(124), 4(123) | 15 | 1(234), 2(134), 3(124), 4(123) | (1234) | 10 | $\emptyset$ | 6 | 4 |
| $n = 5$   | 541    | 1(2345), 2(1345), 3(1245), 4(1235), 5(1234) | 30 | 1(2345), 2(1345), 3(1245), 4(1235), 5(1234) | (12345) | 15 | $\emptyset$ | 24 | 5 |
|           |        | (12)(345), (13)(245), (14)(235), (15)(234), (23)(145), (24)(135), (25)(134), (34)(125), (35)(124) | 12 | (12)(345), (13)(245), (14)(235), (15)(234), (23)(145), (24)(135), (25)(134), (34)(125), (35)(124) | (12345) | 16 | $\emptyset$ | 24 | 5 |

4
\[ |\Omega_2| = T_n = \frac{n(n+1)}{2} = \binom{n+1}{2} = \binom{n+1}{n-1} = (S_{n+1,n}) = \sum_{i=1}^{n} |A_k|_i. \]  

(8)

The cardinality \( |\Omega| \) of the space of forbidden rankings is defined by equation:

\[ |\Omega| = |\Omega_2 \cap \Omega_1| = F_n = 2^n - 1 - n(n+1)/2. \]  

(9)

Evidently, the cardinality \( |\Omega| \) is defined by

\[ |\Omega| = T_n + F_n = 2S_{n,2} + 1 = 2^n - 1. \]  

(10)

It follows from equations (8)–(10) that numbers \( T_n \) can be found among Stirling numbers of the second kind and in Pascal triangle [11, 13].

4. Inranking Generating Sets

Inrankings, as well as any weak orders, are shaped of sets of strict orders. Let us demonstrate this by a simple example.

**Example 3.** Given two strict orders \( \rho_1 = 213 \) and \( \rho_2 = 231 \). Then lying between them (generated by them) weak order is \( \lambda_k = 2(13) \) or in extended notation \( \lambda_k = \{ a_2 \sim a_1 \sim a_3 \} \).

Designate the permutations set of elements of \( A_k \) through \( \Pi = \{ \pi_1, \ldots, \pi_{|A_k|} \} \), and the permutations set of elements of \( \bar{A}_k \) through \( \Pi' = \{ \pi'_1, \ldots, \pi'_{|\bar{A}_k|} \} \).

The set of strict orders \( \Psi_k = \{ \rho_k^g \} \), \( g = 1, \ldots, |A_k| \)!: |\bar{A}_k|!, will be called a generating set for an inranking \( \lambda_k = A_k \succ \bar{A}_k \) if \( \Psi_k = \Pi \times \Pi' = \{ (\pi_u, \pi'_v) \mid \pi_u \in \Pi, \pi'_v \in \Pi' \} \), where the Cartesian product’s element \( (\pi_u, \pi'_v) \) is deemed as a result of concatenation operation \( \rho = \pi_u \pi'_v \) of two permutations \( \pi_u \) and \( \pi'_v \).

**Example 4.** Let \( \lambda_k = (234)(15) \). Then the generating set \( \Psi_k \) consists of \( |\Psi_k| = 3! \cdot 2! = 12 \) following strict orders \( \rho_k^g \), \( g = 1, \ldots, 12: 23415, 23451, 24315, 24351, 32415, 32451, 34215, 34251, 42315, 42351, 43215, 43251 \).

Given cardinality \( |A_k| \), at any \( n \), the number of generating sets is defined by the expression \( (n - |A_k| + 1) \). Hence, the total number of generating strict orders for all possible inrankings can be calculated as

\[ N(A_k) = |A_k| \times |\bar{A}_k| \times (n - |A_k| + 1). \]  

(11)

One can claim that for given \( n \) a union of generating sets for \( |A_k| = 1 \) coincides with the single generating set for \( |A_k| = n! \). All possible cardinalities \( |\Psi_k| = |A_k| \times |\bar{A}_k| \times (n - |A_k| + 1) \) and \( (n - |A_k| + 1) \) for different \( |A_k| \) and \( n = 1, \ldots, 5 \) are shown in the ninth and tenth columns of table 1 correspondingly.

The introduced notations can be illustrated using permutohedra. Figure 3 shows the space of all strict orders for \( n = 5 \) and the space \( \Omega_2 \) of all corresponding inrankings. For clarity, we use five permutohedra which are truncated octahedra with central nodes corresponding to inrankings \( 1(2345), 2(1345), 3(1245), 4(1235) \) and \( 5(1234) \). There are shown only links between inrankings and facet centers of permutohedra containing weak orders. Inranking designations are shown with bold font. Each of the truncated octahedra represents a generating set for inrankings with \( |A_k| = 1 \). In color there are shown subsets of the generating sets belonging to the permutohedron for inrankings at \( |A_k| = 2, 3, 4 \).

Figure 4 shows the generating set for the inranking \( \lambda_k = (234)(15) \) discussed in examples 1–4. One can see that centers of edges and facets of permutohedra correspond to weak orders. As above, there are shown
only links between facet centers of permutohedra and the inranking (234)(15). The number of permutohedra is defined by the cardinality $|A_k| = 3$.

**Figure 3.** The space of all strict orders for $n = 5$ and the space $\Omega_i$ of all corresponding inrankings.

### 5. Conclusion

It is shown that produced by the IF&PA procedure inranking space is a subset of the set of all weak orders with a single symbol of strict order. The cardinality of inranking space is defined by the triangle number for the given number of the discrete elements of the RAV. Knowing properties of the spaces can help to recommend an effective mode of the interval data fusion applications. The properties are particularly important for improvement of search techniques of the consensus ranking for a given preference profile consisting of the inrankings.
Acknowledgments

This work was supported by the Russian Science Foundation, project 18-19-00203, in part of the IF&PA method development, and by the Russian Ministry of Education and Science, basic part of the state task "Science", project 2.5760.2017/8.9, in part of the IF&PA method applications.

References

[1] Hall D L and McMullen S A H 2004 Mathematical Techniques in Multisensor Data Fusion (Norwood: Artech House Publishers)
[2] Meeker W Q, Hahn G J, and Escobar L A 2017 Statistical Intervals: A Guide for Practitioners and Researchers (New York: John Wiley & Sons)
[3] Tofallis C 2009 Least squares percentage regression J. Mod. Appl. Stat. Methods Available at SSRN: https://ssrn.com/abstract=1406472
[4] Brams S J and Fishburn P C 2007 Approval Voting (New York: Springer)
[5] Muravyov S V, Khudonogova L I and Emelyanova E Y 2018 Interval data fusion with preference aggregation Measurement 116 621–630
[6] Kemeny J G 1959 Mathematics without numbers Daedalus 88 571–591
[7] Muravyov S V 2013 Ordinal measurement, preference aggregation and interlaboratory comparisons Measurement 46(8) 2927–35
[8] Muravyov S V and Marinushkina I A 2011 Largest consistent subsets in interlaboratory comparisons: preference aggregation approach Proc. 14th Joint Int. IMEKO TC1+TC7+TC13 Symp. (August 31 – September 2, 2011, Jena, Germany) pp 69–73
[9] Khudonogova L I and Muravyov S V 2018 Interval data fusion with preference aggregation in wireless sensor network: Energy-accuracy trade-off in presence of outliers J. Phys. Conf. Series 1065 072016
[10] García-Lapresta J L and Pérez-Román D 2011 Measuring consensus in weak orders In: Consensual Processes. Studies in Fuzziness and Soft Computing 267 (Berlin Heidelberg: Springer)
[11] Aigner M 1997 Combinatorial Theory (Berlin Heidelberg: Springer-Verlag)
[12] Sloane N J A The On-Line Encyclopedia of Integer Sequences. The OEIS Foundation, https://oeis.org
[13] Anderson J A 2003 Discrete Mathematics with Combinatorics (Upper Saddle River: Prentice Hall)