# RANKINGS: ARE THEY USEFUL FOR MEASUREMENT PRACTITIONERS? ARE THEY IN THE SCOPE OF METROLOGY AND MEASUREMENT SCIENCE? 

Sergey Muravyov<br>Tomsk Polytechnic University, Tomsk, Russia, muravyov@camsam.tpu.ru

Rankings are sometimes considered to be non-empirical, non-objective, low-informative and, in principle, are not worthy to be titled measurement. In our opinion, a ranking is a result of measurement on ordinal scale and is useful to the same extent as any ordinal measurement.

There are a lot of ordinal kind scales in the scope of applied metrology. These are, for example, scales for mineral hardness, earthquake magnitudes, wind force, smell of water, many of scales for different kinds of food quality and many, many others. Point is that measurement results obtained in these scales are frequently treated as some number (score, rank). For example, when measuring hardness on Mohs scale a mineral sample is assigned a number $b$ if it cannot be scratched by standard mineral $b, b=1, \ldots, 10$, and cannot scratch it. This number is, clear, only a label and its use in any additive or multiplicative operation is meaningless.

In fact, the measurement result in the ordinal scale should be the entire ranking of $n$ objects and the ranking is one of points of the weak orders space. In this case there appears a possibility to study a structure of the space, to investigate correlation between rankings and the space cardinality and do many other researches yielding useful information about objects under measurement.

Let us consider our proposal in more details.
Suppose we have $m$ rankings on set $A=\left\{a_{1}, a_{2}, \ldots, a_{n}\right\}$ of $n$ objects. Then we have the relation set $\Lambda=\left\{\lambda_{1}, \lambda_{2}, \ldots\right.$, $\left.\lambda_{m}\right\}$, where each of $m$ rankings (preference relations) $\lambda=\left\{a_{1}\right.$ $\left.\succ a_{2} \succ \ldots \sim a_{s} \sim a_{t} \succ \ldots \sim a_{n}\right\}$ may include $\succ$, a strict preference relation $\pi$, and $\sim$, an equivalence (or indifference) relation $v$, so that $\lambda=\pi \cup v$. Such a relation $\lambda$ is generally called a weak order. The relation set $\Lambda$ can be titled a preference profile for the given $m$ rankings.

For example, let $n=3, m=3$, then we can have three following rankings of objects:

$$
\begin{align*}
& \lambda_{1}: a_{1} \sim a_{3} \succ a_{2} \\
& \lambda_{2}: a_{3} \succ a_{1} \sim a_{2}  \tag{1}\\
& \lambda_{3}: a_{2} \sim a_{3} \succ a_{1}
\end{align*}
$$

They belongs to the space of 13 weak orders as shown in Fig. 1.

Now we can determine a single preference relation that would give an integrative characterization of the objects. Let
a subspace $\Pi$ be a set of all $n$ ! linear (strict) order relations $\succ$ on $A$. Each linear order corresponds to one of permutations of first $n$ natural numbers $\mathbf{N}_{\mathbf{n}}$. We will consider a permutation $\beta \in \Pi$ of the alternatives $a_{1}, \ldots, a_{n}$ to represent the preference profile $\Lambda$ and will call it consensus ranking. It is desirable that, in some sense, $\beta$ would be nearest to the every of rankings $\lambda_{1}, \ldots, \lambda_{m}$.


Fig. 1. The space of all weak orders (rankings) for $n=3$
It is clear that the problem described above is very similar to the problem of voting or group decision where $A$ is a set of $n$ alternatives or candidates which are ranked by group of $m$ individuals (multisensors, voters, experts, focus groups, criteria, etc.).

The ranking $\lambda$ can be represented by an $(n \times n)$ relation matrix $R=\left[r_{i j}\right]$ whose rows and columns are labeled by the objects $a$ and

$$
r_{i j}=\left\{\begin{array}{rl}
1 & \text { if } a_{i} \succ a_{j}  \tag{2}\\
0 & \text { if } a_{i} \sim a_{j} \\
-1 & \text { if } a_{i} \prec a_{j}
\end{array} .\right.
$$

The symmetric difference distance function $d\left(\lambda_{k}, \lambda_{l}\right)$ between two rankings $\lambda_{k}$ and $\lambda_{l}$ is defined by formula

$$
\begin{equation*}
d\left(\lambda_{k}, \lambda_{l}\right)=\sum_{i<j}\left|r_{i j}^{k}-r_{i j}^{l}\right| \tag{3}
\end{equation*}
$$

and may be understood as the number of disagreements between two rankings. Here only elements of the upper triangle submatrix, $r_{i j}, i<j$, of matrix $R$ are summed up. The value of $d\left(\lambda_{1}, \lambda_{2}\right)$ between the first two rankings of our example profile (1) is equal to $1+1+0=2$.

A distance between arbitrary ranking $\lambda$ and profile $\Lambda$ can then be defined as follows:

$$
\begin{equation*}
D(\lambda, \Lambda)=\sum_{k=1}^{m} d\left(\lambda, \lambda_{k}\right)=\sum_{i<j} \sum_{k=1}^{m}\left|r_{i j}^{k}-r_{i j}\right|, \tag{4}
\end{equation*}
$$

From (2), supposing $r_{i j}=1$ for all $i<j$ that corresponds to the natural linear order $a_{1} \succ a_{2} \succ \ldots \succ a_{n}$, it is clear that for any $k=1, \ldots, m$ we have $\left|r_{i j}^{k}-r_{i j}\right|=|1-1|=0$ if $a_{i}^{k} \succ a_{j}^{k}$; $\left|r_{i j}^{k}-r_{i j}\right|=|0-1|=1$ if $a_{i}^{k} \sim a_{j}^{k}$ and $\left|r_{i j}^{k}-r_{i j}\right|=|-1-1|=2$ if $a_{i}^{k} \prec a_{j}^{k}$. Thus, denoting $\left|r_{i j}^{k}-r_{i j}\right|$ through $d_{i j}^{k}$ we have

$$
\begin{gather*}
D(\lambda, \Lambda)=\sum_{i<j} \sum_{k=1}^{m} d_{i j}^{k},  \tag{5}\\
\text { where } d_{i j}^{k}=\left\{\begin{array}{lll}
0 & \text { if } & a_{i}^{k} \succ a_{j}^{k} \\
1 & \text { if } & a_{i}^{k} \sim a_{j}^{k} \\
2 & \text { if } & a_{i}^{k} \prec a_{j}^{k}
\end{array}\right.
\end{gather*}
$$

We can now define an $(n \times n)$ profile matrix $P=\left[p_{i j}\right]$ where

$$
\begin{equation*}
p_{i j}=\sum_{k=1}^{m} d_{i j}^{k}, \quad i, j=1, \ldots, n \tag{7}
\end{equation*}
$$

and the number of voters $m$ of the profile A is present in each of the matrix elements as $1 / 2\left(p_{i j}+p_{j i}\right)=m$, $i, j=1, \ldots, n$. Thus, the value $0.5 p_{i j}$ can be understood as the number of preferences $a_{j}$ over $a_{i}$.

For the example profile (1), we have the following profile matrix $P$ :

$$
\left|p_{i j}\right|=\left[\begin{array}{lll}
0 & 3 & 5  \tag{8}\\
3 & 0 & 5 \\
1 & 1 & 0
\end{array}\right],
$$

where, for instance, $p_{13}=1+2+2=5$.
How to find a single preference relation that would give an integrative characterization of the preference profile A described by matrix $P$ ? Condorcet in 1785 proposed a very natural and now well-known procedure of handling the paired-comparison data contained in the matrix $P$ : in each comparison, the preferred object is the object preferred by a majority of voters, i.e. $a_{i} \succ a_{j}$ if and only if $p_{i j}>p_{j i}$.
However, the binary relation defined Condorcet's rule is not necessarily transitive, i.e. it can be that $a_{i} \succ a_{j}$ and $a_{j} \succ a_{k}$ while $a_{k} \succ a_{i}$. This Condorcet's Paradox of Voting may occur rather frequently: its chances are usually even more than $50 \%$.

Published in 1951 Arrow's Impossibility Theorem has shown that no voting method can satisfy the following three desirable (natural) properties (axioms):
(P1) unanimity (if alternative $a_{i}$ is ranked above $a_{j}$ for all orderings $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{m}$, then $a_{i}$ is ranked higher than $a_{j}$ by $\beta$ ),
(P2) non-dictatorship (there is no $k$-th voter whose preferences always prevail), and
(P3) independence of irrelevant alternatives (for two preference profiles $\Lambda$ and $\Lambda^{\prime}$ such that for all $k$-th voters alternatives $a_{i}$ and $a_{j}$ have the same order in $\Lambda$
and $\Lambda^{\prime}$, alternatives $a_{i}$ and $a_{j}$ have the same order in $\beta(\Lambda)$ and $\beta\left(\Lambda^{\prime}\right)$.
Thus, the Arrow's theorem can serve as a thorough justification of the Condorset's paradox which means that a preference profile is not necessarily transitive even if each $k$-th ranking is a linear order.

In this situation, a reasonable way to get over the paradox is to find such a linear order (permutation) $\beta \in \Pi$ of objects of $A$ that the distance $D(\beta, \Lambda)$ from $\beta$ to the profile $\Lambda$ is minimal, that is

$$
\begin{equation*}
\beta=\arg \min _{\lambda \in \Pi} D(\lambda, \Lambda) . \tag{9}
\end{equation*}
$$

Thus, a solution of the problem (9) is the consensus linear ranking $\beta$ that also is called median order. It should be noticed that the problem may have more than one optimal solution. For our example profile (1) we have $\beta_{1}=\left\{a_{3} \succ a_{1} \succ\right.$ $\left.a_{2}\right\}$ and $\beta_{2}=\left\{a_{3} \succ a_{2} \succ a_{1}\right\}, D(\beta, \Lambda)=5$ (see Fig. 1).

Generally, the space of solutions for the problem (9) is increasing extremely quickly as $n$ rises and it had been proven to be $N P$-hard. However, for reasonable problem sizes (up to $n \approx 50$ ) there are exact algorithms for them to be applied. They typically use branch and bound (B\&B) technique and really can serve as well-defined measurement procedure in ordinal scale.

One can argue that initial rankings are subjective as they are obtained without use of an measuring instrument. Possible answer to this point may be justified with a help of a substitution $m$ voters by $m$ sensors in our problem description. They would vote as people (subjects) but the profile matrix seems to have no distinction of one produced by subjects. And the Condorset's paradox would be getting over by means of the problem (9) solving. Thus, the problem is subject invariant.

Other objection can be that the rankings are nonempirical as they are obtained of a thought experiment and reflect unobservable relations. To comment this position let us remember how reliable are our assumptions when measuring in ratio scale. We believe that the attribute we measure is directly connected to a property under investigation, we think that measurement errors are distributed in accordance to known law, we rely on SI units and so on. However, our security is illusory because the measured attribute can be misleading, error distribution is completely out of our prediction, SI units partition into fundamental and derived ones is only a convention and the measurement instrument is not calibrated as believed, etc. Thus a level of our confidence to classical measurement is arguable. So do rankings, clearly in more considerable extent.

Finally our conclusion is that consensus rankings as they were described above can be treated as ordinal scale measurement results with wide area of application in practical metrology and quality management and control.

